SEMANTIC WAYFINDING FOR INFRASTRUCTURE PROJECTS

Sherif Kinawy, Research Assistant, sheriff.kinawy@mail.utoronto.ca
Tamer El-Diraby, Associate Professor, tamer@ecf.utoronto.ca
Department of Civil Engineering, University of Toronto, Toronto, Canada

ABSTRACT

In the age of Big Data and citizen engagement, access to relevant information on demand is becoming an operational requirement. This trend is also realized in infrastructure projects and will become increasingly important as more municipalities embrace the open data movement. In an effort to utilize the wealth of available data, governments and third parties have developed software systems that facilitate the aggregation of data with wide success along with some limitations. To overcome some of these limitations in the context of infrastructure construction, this paper advocates for the incorporation of semantic wayfinding mechanisms; methods that mine paths through which a user can find information relevant to a subject of interest. This information should also be presented at a respective level of detail and should depend on the profile of each user which can be extracted explicitly or implicitly. The benefits of this form of content filtering can be realized through facilitating the navigation of technical documents by the public and to a certain degree, the other way through examining community-generated content by project staff.

This paper reviews a number of wayfinding techniques and emphasizes the need for incorporating explicit and implicit user profiles in the selective application of these techniques. A synthesis of the techniques is provided through an examination of their assumptions and limitations. Finally, the paper will present examples of cases in which this application of wayfinding in infrastructure projects can enhance the experience of planners, designers, engineers and members of the community.

Keywords: semantic wayfinding, community engagement, document visualization, information enrichment

1. INTRODUCTION

The open data movement has given rise to a wealth of information along with a growth in user expectations. This movement has been supported by a tangible improvement in access to information through online channels, recommendation systems, and online social networks (Ferro et al. 2013). This availability of data is promising by many measures (Linders 2012, Johnston and Hansen 2011) since it can allow for better planning through better monitoring. For instance, web platforms such as SeeClickFix utilize standardized data structures to crowdsource non-emergency maintenance issues around cities through online and mobile applications. Standards such as the Open311 API have facilitated the production of community-driven solutions. While these patterns mark progress for advocates of transparency and open data, there are issues of quality and quantity following a trend that continues as pointed out in the past by Kautz et al. (1997).

Nevertheless, engaged citizens require access to relevant information on demand, and this operational requirement is also realized in infrastructure projects; a trend that will become increasingly important as more municipalities embrace the open data movement. In an effort to utilize the wealth of

available data, governments and third parties have developed software systems that facilitate the aggregation of data with wide success along with some limitations. To overcome some of these limitations in the context of infrastructure construction, this paper advocates for the incorporation of semantic wayfinding mechanisms; methods that mine paths through which a user can find information relevant to a subject of interest. This information should also be presented at a respective level of detail, and should depend on the profile of each user which can be explicitly or implicitly extracted. The benefits of this form of content filtering can be realized in the way it would facilitate the navigation of technical documents by the public and to a certain degree, the examination of community-generated content by project staff.

This paper investigates semantic wayfinding as a tool for document enrichment through reviewing a number of wayfinding techniques. The article also emphasizes the need for incorporating explicit and implicit user profiles in the selective application of these techniques. Finally, the paper will present examples of cases in which this application of wayfinding in infrastructure projects can enhance the experience of planners, designers, engineers and members of the community.

2. RESEARCH METHOD

This paper follows a method typical of a review paper with three primary elements: a literature review, comparison of techniques and a proposed solution to address opportunities for improvement. The paper also suggests a method for validation in the form of a proposal for a software system. The literature review was conducted by identifying the need for automated information filtering, and investigating wayfinding techniques and their limitations.

The review starts with a problem definition, a description of the associated knowledge encapsulated in this process in the form of an ontology for infrastructure planning and communication. This ontology represents the knowledge that encapsulates an important question: how do the public look for information on projects and components in their cities? The primary concepts that are used to answer this question are the Actor and Profile. These concepts also constitute the main contribution of this work to existing theories. The review also takes into consideration that every infrastructure project has characteristics that are defined by its components, technologies associated with it and its perception by the public. Such characteristics, referred to as project profiles, can be generalized as a member of the ontological concept, Project Profile. On a much larger scale and in a similar way, cities have a City Profile which determines how information about a city affects citizen preferences. Evidently, profiles for a city or a project can be readily available to planners in a participatory planning environment. The elements of the ontology relevant to this application are further demonstrated through the proposed software system.

3. REVIEW OF THE LITERATURE

The ways by which internet users search for information are complex. This complexity creates an environment in which realizing patterns for research and application is a difficult task. On the theoretical front, scholars have attempted to identify distinguishable patterns based on user. Over the past 20 years, several research projects tackled the task of characterizing web surfing patterns. Huberman et al. (1998) identified web surfing patterns such as “the balkanization of the web structure” which led to a difficulty in finding information despite succeeding in finding regularity in the browsing trails that can be modeled. Yet, Spink et al. (2000) affirm that at that time, we still had little understanding of how the public consumed the information they accessed on the web. By the start of the last decade, significant statistics were available but formulating patterns was still lacking. For example, Cockburn and McKenzie (2001), in a study on a limited number of users, state that 81% of pages are revisited. Moreover, they noted that their subjects only visited each of these web pages for a short period of time (2001). This observation shows that users attempt to take in a lot of information in a short period of time; hence affecting their propagation behavior from one page to another. Users also have to traverse through a large number of

pages to get to the information they need. Nevertheless, scholars assert the lack of commonality (Henzinger and Lawrence 2004, Cockburn and McKenzie 2001) in user behavior.

The social web has complicated the nature of web browsing and sharing of information in the form of news or social recommendations. It also brought user profiles and influence networks into the picture as an important feature of social influence. User profiles now play a greater and more prominent role in how users consume online information (Abel et al. 2011, Sugiyama et al. 2004, Kautz et al., 1997). Essentially, users have different profiles on the web and this establishes these profiles as potential inputs in analyzing browsing behavior and predicting suitable paths of action and suitable communication methods. Nevertheless, the social web also gave rise to an increasing number of theories on wayfinding (Dinet and Kitajima 2011, Hochmair and Raubal 2002, Card et al. 2001); an attempt to capitalize on seemingly chaotic behavior that was even more magnified through social media. One particular conclusion that garnered consensus is that there is an overwhelming amount of information being generated and communicated at an unprecedented rate.

On the practical side, e-commerce and marketing applications operate in this new era under the premise that online communities depend on trust; an important concept in e-loyalty (Reichheld and Schefter 2000). While trust is an important concept from a practical perspective, it is only one of many factors that affect how users decide to consume information. Additionally, e-commerce operates towards a target user group which limits the transferability of the techniques used in most e-commerce systems (Romano and Fjermestad 2001) to other less commercial fields. This limited application is also exhibited in the focus of e-commerce, in many cases, on retaining existing clients rather than attracting new customers, as suggested by Reichheld and Schefter (2000).

Theoretical platforms attempted to tackle the subject of wayfinding as a research problem on a grander scale, taking an analogy from the physical world where maps are used for navigating space and orienting users (Hochmair and Raubal 2002). In the built environment, humans have become adept at utilizing several tools that depend on signage to identify suitable paths or ways and mark them for others. This document takes this accumulated knowledge into consideration despite the disparity in context between navigating information paths and physically navigating the built environment. Yet navigation and browsing share common elements such as creating paths based on the interests of the user and purpose of navigation.

While browsing itself may not be defined as a problem, supported wayfinding presents an efficient method that enhances the process through which users can find relevant information, and more importantly, make sense of this information in a short time (O'Day and Jeffries 1993). In the following section, a discussion of various techniques is presented as an example of the plethora of algorithms used to guide online wayfinding.

4. INVESTIGATING WAYFINDING

The use of wayfinding extends the benefits of web search. Wayfinding in a more general context can facilitate benefits to specific disciplines that are now increasingly relying on the web for communication and information dissemination. The review of the literature resulted in the identification of a number of algorithms and techniques that can potentially enhance the user experience for members of the public exploring various infrastructure projects. This comparative study investigates the suitability of specific wayfinding algorithms to the narrower context of infrastructure projects. The aim of this investigation is to:

a) study infrastructure/planning-specific wayfinding issues, their unique attributes and techniques that fit some specific needs of the domain;

b) extend the wayfinding agenda to project engineers and planners, and help them navigate technical and community related documents to help improve their designs given the overwhelming number of documents and studies available to each; and

c) enrich wayfinding algorithms with an ontology to support knowledge integrity.

The wayfinding techniques identified in the review can be grouped into three categories: 1) trail-based for navigating a set of documents; 2) concept-based for navigating information within a document or a number of documents; 3) prediction-based for navigating different forms of information by type or behavior.

4.1. Navigating A Set Of Documents

Given the fast pace of web navigation, document/page hopping comprises a large portion of web browsing where users are constantly deciding where they should go next. Traditionally, information retrieval algorithms were used to predict this behavior, aid in browsing and capitalize on commercial marketing opportunities. More recent research follows a theoretical framework set by Bates (1989) called Berrypicking in an attempt to better mimic user behavior which can seem random at times. The trails left behind by users as they navigate from one page to another has also been further studied and consequently gave rise to concepts such as scent trails as discussed by Olston and Chi (2003) and Chi et al. (2001). More recently, search trails and subtrails were studied as a method to inform queries by White and Huang (2010) and Singla et al. (2010) through browser logs, and West and Leskovec (2012) through a Wikipedia-based game.

4.2. Navigating information within a document

In specific cases where a predefined set of information is available, a user can perform a set of queries on this limited information within one or more documents. Teevan et al. (2004) explains how user behavior can be used to rank search results using specialized algorithms. This application is more focused on specific searches instead of page hopping and trail setting. This form of implicit user profiling was also used by Shen et al. (2005) to improve information retrieval efficiency.

4.3. Navigating different forms of information

The former two categories have been proven to work well for structured data of a homogonous nature, for example, a set of sports news articles. This simplicity offers room for improvement by taking into account social influence networks together with comments from users and automated behavior recognition. Adamic and Adar (2005) address search techniques in the case of unstructured, semi-structured or incomplete data which resort to an analysis of social networks. Prior to the proliferation of social networks in their current form, social networks and small worlds were proposed as a useful information navigation tool for communities of interest (Killworth et al. 2006, Liben-Nowell et al. 2005, Kleinberg 2000).

5. SOLUTION

The benefits offered by wayfinding can be weighed down by the overwhelming increase in the amount data and complexity of user navigation. A number of techniques were identified along with some limitations which include the absence of context and a knowledge base. However, a combination of similar ontology-based techniques and infrastructure-specific context-based solutions would offer a more customized experience for each user. The solution provided in this paper depends on three main concepts: user profiling, visualization, and a knowledge base.

Infrastructure planning and construction is unique in many ways. In the context of urban centres, such projects and their respective stakeholders have the following characteristics:

1) Long-term, dynamic project plans
2) Involvement of schedules, cost sheets, drawings and visuals
3) Stakeholders vary from very technical to policy-savvy to the general public
4) Can be viewed as a set of components, functions and impacts, or all at the same time
5) Different parameters depending on context: location, time, and conditions

To enable effective wayfinding with added context, this paper proposes that the user profile should be an integral part of the system. Furthermore, a knowledge structure is integrated to enhance the integrity of the wayfinding mechanism. Visualization is an emerging yet common concept that is best explained using a hypothetical example: a municipal project that involves planning a bicycle lane along a main arterial. The project documents for this example include the environmental assessment, route maps, and cross-sections of various intersections.

5.1. The User Profile

Traditionally, users seeking information beyond the default document order were required to manually navigate list of documents or surf documents through cross-links. More advanced users execute search queries to jump to specific pieces of information or relevant information about a project.

While identifying relevant media or online documents has been studied using user profiles Bilenko, M. & White, R.W. (2008, Razmerita et al. (2003) demonstrated that modeling user profiles can be greatly enhanced using an ontology. A typical user profile encapsulates identifying basic demographic information. Advanced profiles can track user interests using browsing history to include relevant topics. This project goes beyond typical user attributes to add parameters that reflect browsing preferences specific to infrastructure projects.

In the case of the bicycle lane project, for example, some users may be interested in the size of the bicycle lanes, the pavement material and intersection types. Other users can be more interested in the impacts of construction in terms of noise and vibrations reaching their front yards. Finally, some users may be cyclists who are more interested in the function these lanes serve in terms of improving their commute time. To account for these differences in interests for the same project, one primary modality for user profiles classifies users into either function-oriented, impact-oriented, or product-oriented. A product-oriented classification, typically encountered in engineering documents, is shown in Figure 1.

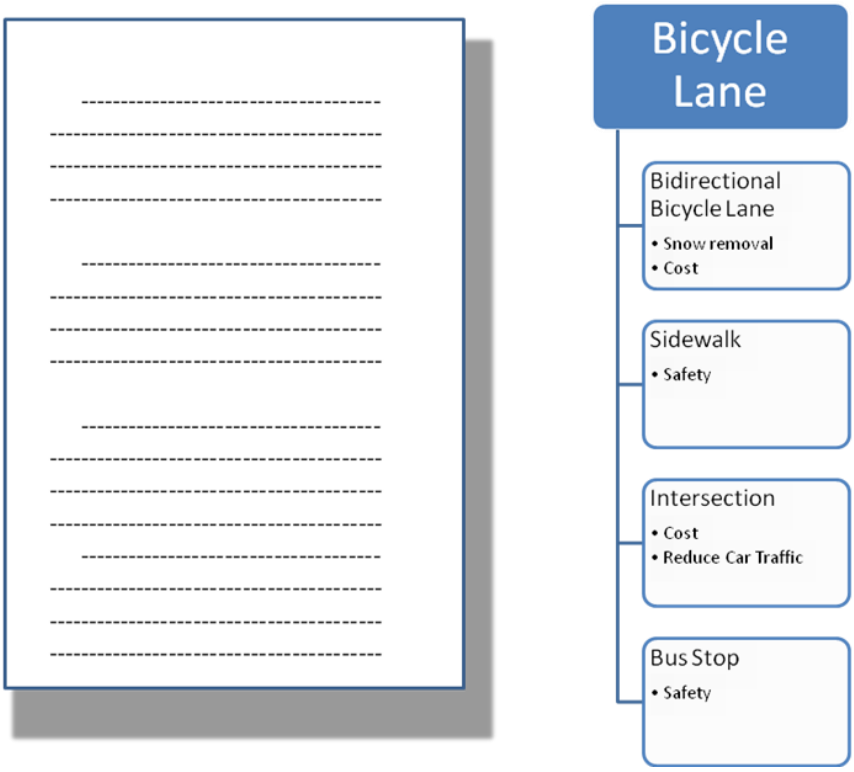

Figure 1: Classifying infrastructure concepts based on project documents

5.2. Visualization

Lately, advancements in Geographic Information Systems, interactive mapping techniques and other visualization technologies have enabled more creative ways of presenting information. Visualization now plays a more important role in communication and exchange. In the case of the bicycle lane project, visualization is essential in highlighting connections and important concepts, and providing a rapid response through dynamic visual structures through space and time as shown in Figure 2.

5.3. Knowledge Structure

The knowledge structure is formalized through an ontology that represents the primary concepts involved in the communication process as well as components of infrastructure projects in general. In doing so, the ontology links the various user profiles identified earlier with communication preferences and various products, functions and impacts. The pivotal concepts in this sub-ontology are shown in Table 1 along with a selection of attributes.

Table 1: Some ontological concepts

Concept	**Attribute**
City Profile	Climate (harsh winters, rainy summers)
	Major events (Summer Olympics, marathon, technology exhibition, arts festival)
	Political atmosphere (stable, election year)
	Governance structure (elected mayor, partial federal involvement)
	Demographics (age distribution, unemployment)
	Type of economy (service-oriented economy, industrial economy)
	Types of residing communities (student community, bedroom community)
	Social attitudes (welcoming to news investments, resistant to change)
Project Profile	Duration (short or long term, temporary or permanent)
	Common controversial elements (road closures, noise)
	Functions and impacts (safety, environment)
	Location-related characteristics
User Profile	Function-oriented, impact-oriented, product-oriented
	Daily commute times
	Family size
	Pet ownership
	Daily activities (shopping, recreational, hobbies,
	Main modes of transportation (bicycle, public transit, automobile)
	Special needs (kids park, accessibility)
	Attitudes towards past projects (supportive, advocate)

City Profile: A collection of attributes including climate, major events, political atmosphere, governance structure, demographics, types of economy and communities and social attitudes.

Project Profile: Properties of projects can be specified within this class. Such properties can include duration, common controversial elements, functions and impacts, and location-related characteristics. The project profile can inherit some attributes from the city where it is located.

User Profile: Based on the type of profile (function-oriented, impact-oriented, product-oriented), different parameters are included for each profile. For instance, impact-oriented profiles would include information on daily commute times, family size, pet ownership, daily activities, main modes of transportation, special needs, and attitudes towards past projects.

Through the resulting configuration, the structure in Figure 1 can be transformed by adding visualization to the more dynamic structure in Figure 2. Using the ontology as a core, a reasoning engine can automatically distill information relevant to the user's profile based on whether they are function-oriented, impact-oriented, or product-oriented. The reasoning engine can also display relevant links where products, functions and impacts are linked directly to relevant text documents, images or videos for a more complete picture.

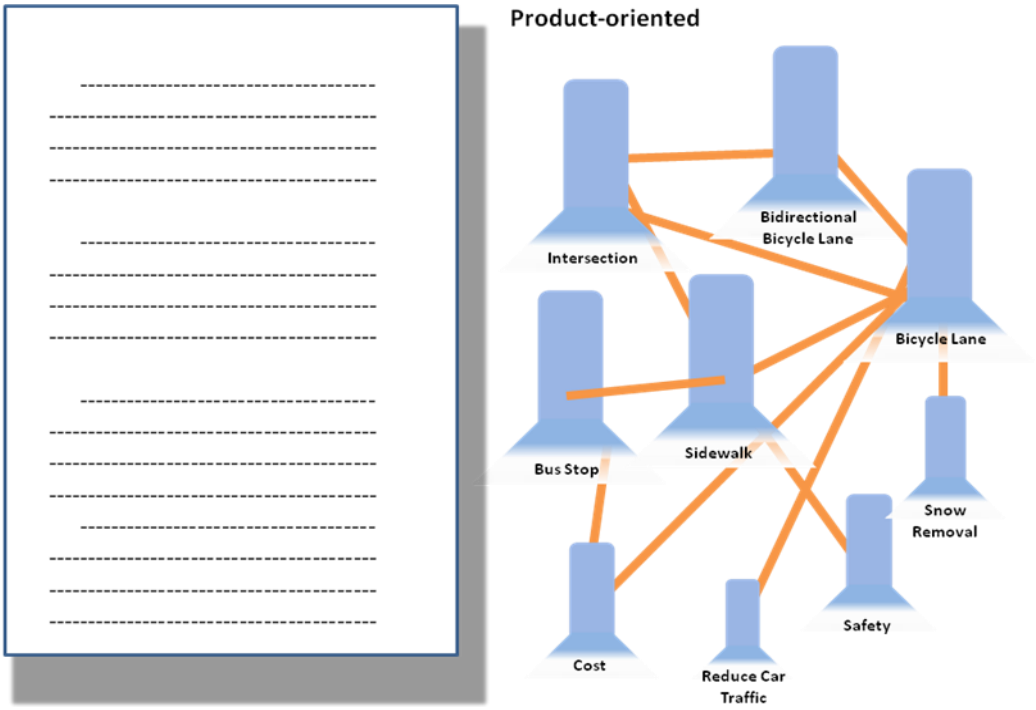

Figure 2: Visualizing infrastructure concepts based on project documents and user profiles

6. DISCUSSION AND CONCLUSIONS

The proposed system uses a combination of visualization, user profiles and knowledge management to enhance the user experience in information systems. Based on the context generated from the system, one or more of the wayfinding techniques identified in the literature review can be used to guide the user through a ranked and visually enhanced choice of concepts and related files. These files can be in various media forms according to the user's profile whether that is text-based, illustrated images or maps, or annotated videos.

While this paper focused on the process of user participation, the proposed solution promises to impact the outcome of the process through enhancements to the decisions. These improvements in the ability to navigate complicated technical documents and public feedback empower stakeholders to make more informed decisions. The software specifications presented in this paper form the basis for a validation framework for the proposed user-specific, knowledge-enabled wayfinding system. Future steps

will include developing a prototype software package for testing and benchmarking using predefined infrastructure-engineering-specific scenarios.

Together with the knowledge component, sorted content and the appropriate algorithm, this project can contribute to the efficiency of the process by which members of the public can find information about a specific infrastructure project. Towards this main goal, the following functions were attained: 1) customizing project news feeds for each user, and 2) filtering through information that is relevant to the user's profile and reducing the search space.

REFERENCES

Abel, F., Gao, Q., Houben, G. J., & Tao, K. (2011). Analyzing temporal dynamics in twitter profiles for personalized recommendations in the social web.

Adamic, L. and Adar, E.. (2005). How to search a social network. *Social Networks*, 27(3), 187–203.

Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407–424.

Bilenko, M. & White, R.W. (2008). Mining the search trails of surfing crowds: Identifying relevant websites from user activity. In *WWW*.

Card, S. K., Pirolli, P., Van Der Wege, M., Morrison, J. B., Reeder, R. W., Schraedley, P. K., & Boshart, J. (2001, March). Information scent as a driver of Web behavior graphs: results of a protocol analysis method for Web usability. In Proceedings of the *SIGCHI conference on Human factors in computing systems* (pp. 498-505). ACM.

Chi,E.H., Pirolli, P., Chen, K., & Pitkow, J. (2001). Using information scent to model user information needs and actions and the Web. In CHI.

Cockburn, A., & McKenzie, B. (2001). What do Web users do? An empirical analysis of Web use. *International Journal of human-computer studies*, 54(6), 903-922.

Dinet, J., & Kitajima, M. (2011). Draw me the Web: impact of mental model of the web on information search performance of young users. In *23rd French Speaking Conference on Human-Computer Interaction* (p. 3). ACM.

Ferro, E., Loukis, E. N., Charalabidis, Y., & Osella, M. (2013). Evaluating Advanced Forms of Social Media Use in Government.

Henzinger M. & Lawrence, S. (April 2004). Extracting knowledge from the *World Wide Web*. *PNAS*, 101. Retrieved from http://www.pnas.org/cgi/reprint/101/suppl_1/5186.

Hochmair, H., & Raubal, M. (2002). Topologic and metric decision criteria for wayfinding in the real world and the WWW. Proceedings of *Spatial Data Handling (SDH'02)*, Ottawa.

Huberman, B.A., Pirolli, P., Pitnow, J.E., & Lukose, R.M. (1998). Strong regularities in *World Wide Web surfing*. *Science*, 280(5360), 95– 97.

Johnston, E. W., & Hansen, D. L. (2011). Design lessons for smart governance infrastructures. *American Governance*, 3.

Kautz, H., Selman, B., & Shah, M. (1997). Referral Web: combining social networks and collaborative filtering. *Communications of the ACM*, 40(3), 63-65.

Killworth, P., McCarty, C., Bernard, H., & House. M. (2006). The accuracy of small world chains in social networks. *Social Networks*, 28(1), 85–96.

Kleinberg, J.M. (2000). Navigation in a small world. *Nature*, 406(6798), 845–845.

Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., & Tomkins, A. (2005). Geographic routing in social networks. *PNAS*, 102(33), 11623–11628.

Linders, D. (2012). From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. Government Information Quarterly.

O'Day, V.L. & Jeffries, R. (1993). Orienteering in an information landscape: How information seekers get from here to there. In *CHI*.

Olston, C. & Chi, E.H. (2003). ScentTrails: Integrating browsing and searching on the Web. *TCHI*, 10(3),177–197.

Razmerita, L., Angehrn, A., & Maedche, A. (2003). Ontology-based user modeling for knowledge management systems. In User Modeling 2003 (pp. 213-217). Springer Berlin Heidelberg.
Reichheld, F, & Schefter, P. (July-August 2000). E-loyalty: Your Secret Weapon on the Web. Harvard Business Review (R00410).
Singla, A., White, R.W., & Huang, J. (2010). Studying trailfinding algorithms for enhanced web search. In *SIGIR*.
Spink, A., Wolfram, D., Jansen, M.B.J., & Saracevic, T. (2000). Searching the web: The public and their queries. Journal of the American Society for Information Science and Technology, 52(3).
Sugiyama, K., Hatano, K., & Yoshikawa, M. (2004). Adaptive web search bassed on user profile constructed without any effort from users. In Proceedings of the 13th international conference on World Wide Web (pp. 675-684). ACM.
Teevan, J., Alvarado, C., Ackerman, M.S., & Karger, D.R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. In *CHI*.
West, R., & Leskovec, J. (2012). Human Wayfinding in Information Networks. Web User Behavioral Analysis and Modeling, www2012, France.
White, R. W. & Huang, J. (2010). Assessing the scenic route: Measuring the value of search trails in Web logs. In *SIGIR*.