

Empirical Studies of the Searching Behavior of Novice Users when Looking for Technical Information

Fernando Figueira Filho^{*}
Institute of Computing
State University of
Campinas (Unicamp),
Brazil
fernando@las.ic.unicamp.br

Gary M. Olson
Bren School of Information
and Computer Sciences
University of California,
Irvine, USA
gary.olson@uci.edu

ABSTRACT

In open and distributed environments such as web forums, information is produced at an enormous rate by a large number of users with different backgrounds. Thus, effectively communicating information-seeking needs in the form of keywords and filtering information produced by other community members still remain an issue. With the aim of addressing this issue, a community search engine is being developed. We will use this system to conduct a series of studies with a group of novice Linux users. Experiments will be held at University of California, Irvine (UCI), University of Campinas (Unicamp) and University of São Paulo (USP).

ACM Classification Keywords

H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces—*web-based interaction, collaborative computing, organizational design*

General Terms

Human factors

Author Keywords

user study, information seeking

OVERVIEW

In recent years, we noticed a paradigm change in the World Wide Web. New web application functionalities incite a growing change in the role of a user from a mere information consumer to an active knowledge producer. Indeed, user participation in knowledge-sharing activities is now considered an important social process. Web logs (blogs), problem-solving discussion forums, wiki-based tools are a few examples of a new class of knowledge-sharing applications which

^{*}Work done while at the University of California, Irvine under a grant from CAPES.

have been considered as a cost-effective platform for collaborative knowledge production. Although we observe an increase in the amount of information that is produced collaboratively, a closer look shows that the scope of the collaborative activity has never been so fragmented. Online communities interact obliviously to the geographical barriers that separate their members. Bannon and Bødker [1] defines the web information space as “one of the most open—in the sense of accessible—electronic spaces that exists, while at the same time be one of the most closed—in the sense that due to the heterogeneity of users and possible use situations, the possible interpretations of the information that is presented is impossible to know”. Furthermore, the design of knowledge-sharing applications plays an important role in the the characterization of online communities in the Web.

In this work, we consider the open-source community and their practices on supporting members to solve technical problems online. Web forums are important repositories of information and also a means to articulate distributed software development teams. But as far as the design of knowledge-sharing applications is concerned, there are some issues regarding searching and knowledge reuse. In the particular case of problem-solving tasks, sometimes one needs information that may be scattered throughout a series of different postings within the web forum. As a result, searching systems should be able to support users to find, assess and exploit many distinct threaded discussions in order to solve a particular problem. However, the majority of these systems are exclusively based on the keyword-search approach, i.e. a text box with a search button. The problem with this approach is that it assumes that the user knows how to formulate the search query. This assumption also has some other implications when considering the online community and their collaborative activities. On one hand, users are part of wider social structures that have different classification practices [2]. For instance, people annotate threaded discussions differently and perceive their activities from distinct, situated perspectives. On the other hand, there are also discrepancies in the level of expertise between community members. This issue is particularly relevant when exploring information in knowledge domains within which the user does not have a strong background, e.g. a Linux novice user searching for problem solutions in a web forum. Experienced users may use very specific terms to categorize content and describe

solutions. Our hypothesis is that the technical expertise gap between novice and experienced users may lead to several mismatches and difficulties when both types of users are communicating using the web forum system.

With the aim of addressing those issues, a community search system is currently under development and will be evaluated by conducting user studies at UCI, Unicamp and USP. The studies will be divided into three phases. First, participants will be recruited to enroll in individual usability sessions. The aim is to identify the strategy used by users when performing search tasks on technical topics and the reasoning behind their actions. We are particularly interested in better understanding the novice users' main difficulties when formulating queries and identifying relevant results. The goal of the second phase is to gather assessment data about the effectiveness of our search engine, in comparison with the conventional keyword-based approach. These results will serve as a basis for determining the appropriate sample sizes and refining the procedures and materials for the third phase, which aims to validate our search system.

BENEFITS FROM ATTENDING

The purpose of attending the doctoral colloquium is to come up with new ideas about how to identify and analyze the impacts of our tool in this collaborative environment, i.e. the web forum community. We want to extend the concept of search efficiency to a broader definition, which will be capable to reflect not only the relationship between a single user and the search tool, but also between a given online community and the tool. Since the Web became an important platform for collaboration and information sharing, we want to better understand if the design decisions made upon an artifact, i.e. a search engine, are capable to engender a stronger sense of community within a distributed, loosely-coupled group of users and, at the same time, affect the way these users are collaborating within the online community. Additionally, issues such as vocabulary naturalization [2], membership and identity [3] still need to be addressed in the context of web forums. We believe that CSCW researchers can contribute with new insights about methods, e.g. how to address those issues, given the virtuality of the online community and their distributed collaborative practices.

RESEARCH CONTRIBUTIONS

The design of search systems with the purpose of supporting the informational requirements of different users poses a considerable challenge. The challenge is to cope with semantics considering a common information space, i.e. the web forum, in which information is shared, classified and consumed in multiple social contexts at once. On one hand, web forum users are members of multiple communities of practice [4], thus employing a different vocabulary and distinct levels of specificity for classification. On the other hand, in the particular case of web information spaces, each user can play two distinct roles—namely, information producer or consumer—which in practice constitute two different and non-simultaneous situated activities.

The purpose of this research is twofold: (a) to explore the

main difficulties faced by novice Linux users when searching on web forums for support information and reusing knowledge produced by other community members and (b) drawing upon the concept of communities of practice, analyze the impacts of our search system in the online community, i.e. how an artifact—the search system—can engender the legitimate peripheral participation [3], strengthening the online community. As a result, we expect to address two issues. First, the majority of today's researchers tend to evaluate search systems focusing only on measuring ranking quality with variants of precision and recall, which is called Cranfield or "batch mode" style of evaluation. A consistent criticism of the Cranfield style of evaluation is that it does not reflect the wide range of user behavior associated with information seeking activities. We expect to extend the understanding of individual search behavior, giving emphasis to the collaborative activities that take place within the web forum and analyzing patterns of group behavior, i.e. who contributes to who and how this could help us to design a better search system. Second, existing user studies of novices' behavior tend to analyze situations in which users are searching on general topics, usually in a non-collaborative environment. In this study, we are focusing in the information exchange about technical topics. We believe that, when confronted with searching situations in unfamiliar knowledge domains, novices can provide useful information about their strategies to interpret information produced by other community members. Based on the findings provided by the empirical study we will conduct, we expect to develop an alternative to the conventional keyword-based approach, which can assist novice users to look for information in a straightforward manner and also provide a more efficient way to reuse the knowledge produced in online communities.

ADDITIONAL AUTHORS

João Porto de Albuquerque
School of Arts, Sciences and Humanities, University of Sao Paulo, Brazil, joao.porto@usp.br

Paulo Lício de Geus
Institute of Computing, State University of Campinas (Unicamp), Brazil, paulo@las.ic.unicamp.br

REFERENCES

1. L. Bannon and S. Bødker. Constructing common information spaces. In *ECSCW'97: Proceedings of the fifth European Conference on Computer-Supported Cooperative Work*, pages 81–96, Norwell, MA, USA, 1997. Kluwer Academic Publishers.
2. G. Bowker and S. Star. *Sorting Things Out: Classification and Its Consequences*. MIT Press, Cambridge, MA, 1999.
3. J. Lave and E. Wenger. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, 1991.
4. E. Wenger. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, 1999.