Challenges in anticipatory information management under network constraints

Felipe Meneguzzi
Jean Oh
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213, USA
meneguzz@cs.cmu.edu
jeanoh@cs.cmu.edu

Katia Sycara
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213, USA
katia@cs.cmu.edu

Simon Parsons
Brooklyn College,
City University of New York
Brooklyn, NY 11210, USA
parsons@sci.brooklyn.cuny.edu

Timothy J Norman
Dept. of Computing Science
University of Aberdeen
Aberdeen, AB24 3UE, UK
t.j.norman@abdn.ac.uk

***Abstract*—As various militaries move towards a vision of Network Centric Warfare and Network Centric Operations, multiple issues regarding information acquisition and presentation arise as one tries to use the full potential of a networked environment while overcoming the limitations imposed by the realities faced by the warfighter. In such an environment, a user processes, communicates, and shares information electronically, creating a demand for intelligent software agents for information gathering and management. When a user must make critical decisions under time pressure, such an agent can anticipate the user's information needs, prefetch and present information to the user in an appropriate format at the right time. In this paper, we discuss some of the key challenges in designing unobtrusive information agents, and suggest directions for further research in order to overcome or mitigate these challenges.** [1]

I. Introduction

In complex, dynamic environments, such as those faced by a warfighter during combat, one must process, communicate, and share information electronically potentially leading to the kind of *information overload* observed over the internet [3], creating a demand for intelligent software agents for information gathering and management. Like a human assistant, a software assistant shares some parts of the work environment with the user, and may have access to the user's private information such as the history of information queries, email messages or calendar events. Thus, when the user must make critical decisions under a time-pressured situation such an intelligent agent may anticipate the user's information needs, prefetch information, and present it to the user in an appropriate format at the right time.

In this paper, we discuss the main challenges in designing the information gathering mechanism for intelligent information agents arising from network limitations. We start the paper with an overview of proactive information assistance in Section II within which we set the context for information gathering, and proceed to introduce information gathering in Section III while discussing the key challenges in Section IV. Finally, we suggest directions for research to overcome the challenges faced by a proactive information agent in meeting a user's information needs, concluding the paper in Section V.

[1]Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

II. Proactive Information Assistance

In order to tackle the huge volume of information available on the internet and to compensate the delay associated with retrieving information in low bandwidth settings, internet browsers and search engines attempted to deploy several prefetching schemes [1], e.g., Mozilla Link Prefetching and Google Web Accelerator. These approaches took advantage of the relations between documents implicit in the hyperlinks connecting them with the assumption that users are likely to need documents related to recently consulted ones. Nevertheless, prefetching in web browsing has not seen widespread adoption due to several shortcomings [4]. In particular, the accuracy for predicting pages actually requested by a user is usually not high to justify the additional drain on communication resources, which was believed to be consequence of the user-independent method of inferring future needs [4]. Recent work on anticipatory information gathering [5] has shown that the relevance of the presented information can be greatly enhanced by using a personalized model of user activity that includes monitoring the user and predicting the information most likely to be needed by that user [2].The models generally employed by these information assistants to infer a user's information needs are generally based on a mixture of stochastic planning planning and bayesian inference [6], such as the model used in the CALO assistant [7]. One of the main assumptions in the use of these models is that the assistant can take immediate action as soon as more of the user's needs are discovered. This assumption, however, is not necessarily accurate when one considers users operating in inhospitable environments with unreliable communications.

III. Incremental Information Gathering

As an agent collects new observations and updates its prediction of the user's plan, a set of information needs relevant to the predicted plan is inferred and is passed to the information gathering component, which generates an information plan to meet these needs. As a user's information needs dynamically change, the information gathering plan must be revised in order to satisfy the new demands. This information plan needs to be revised whenever information needs change. The plan can be rewritten from scratch to accommodate the new needs, discarding any existing plan. However, such a naïve approach makes sense only in problems where the information demands are more or less randomized such that little of existing plan can be reused. However, in more structured problems, information needs gradually change over time, an incremental approach is more suitable that incorporates new demands into an existing plan only by modifying small parts that have been affected by the new demands instead of re-solving the entire problem. A new demand may indicate changes in the properties (*e.g.* priorities) of information needs in the current plan, or include a new set of information needs. Note that a new demand may arrive during the middle of plan execution; that is, the agent may be already retrieving information from remote sources. An information need may become obsolete if it is no longer in the current demand. In this case, the agent may cancel the retrieval of the information completely, but the agent may instead keep it with a lower priority just in case it will be needed again. Here, our research focus is on developing good strategies for such decisions so that the agent maximizes the assistance while it minimizes the use of computing and networking resources that the agent shares with the user.

IV. Challenges in information management

Even with an improved model of future user needs, key challenges still remain. Some users are concerned that prefetching may violate security constraints; such as by retrieving illegal material; for users who pay for network usage based on the volume of data retrieved, unnecessary prefetching adds extra cost. Lastly, prefetching activities compete with the user for computational resources and network usage, and if multiple users share a network link, prefetching can quickly overwhelm the available communication channels.

Within battlefield scenarios, besides the fast-pace of change of information needs, proactive assistance has to deal with additional challenges to the process of information gathering.In particular, the challenges of delivering information to a warfighter can be divided into two categories: *i)* device limitations; and *ii)* communications limitations. Device limitations refer to the ability of a computing device, especially portable computing devices such as PDAs, to retrieve, store, process and display information. Communications limitations refer to the ability of a computing device to discover and effectively use available communications channels to its non-local information sources.

While mobile computing devices have improved significantly in recent years, battery technology does not improve at the same pace, and in order to maintain battery life, portable electronics necessarily provide quantitatively limited capabilities when compared to fixed computing devices. Thus, mobile computers cannot store and process all the possible information relevant to a particular domain. Tradeoffs are often necessary, as exemplified by the way Google's mapping service operates on mobile devices, with very little information being cached in device itself, making it completely depending on a network or cellular connection. In a battlefield environment, communication capabilities can be severely limited or completely negated by a number of factors, including distance to high-powered communication hubs, electronic *noise* coming from the multitude of communication and jamming devices present in modern military vehicles, or the unavailability of long-range communications equipment. Thus, information gathering agents must have strategies to cope with sudden degradation of communications capabilities, while maintaining a modicum of assistance capabilities.

V. Discussion

Degradation of communication capabilities can arise due to many factors and be manifested in multiple ways. Our current research focuses on developing strategies to provide information assistance under four communication situations: *i)* no communications; *ii)* brittle communications on a dedicated channel; *iii)* brittle communications on a shared channel; and *iv)* full communications. For each of these situations, certain strategies are more appropriate than others. If there is the expectation of no communications whatsoever during a mission, information can be prefetched before the mission starts, and so the task of an information agent is to predict likely critical scenarios during a mission and cache information before the mission starts in order to provide it during the mission. If communications are brittle within a dedicated communication channel, then an agent should try to maximize the amount of information that is gathered during the temporal windows in which information is available, whereas if the communication channel is shared, an agent must negotiate some compromise with other users sharing the channel in order not to jeopardize the mission of others in its vicinity. Finally, if communications are fully available, then information gathering can be delayed to the last moment, but strategies are needed not to overload the user with irrelevant information.

References

[1] Virgilio Almeida, Azer Bestavros, Mark Crovella, and Adriana de Oliveira. Characterizing reference locality in the www. Technical report, Boston, MA, USA, 1996.

[2] G. Barish and C. A. Knoblock. Speculative plan execution for information gathering. *Artif. Intell.*, 172(4-5):413–453, 2008.

[3] Pattie Maes. Agents that reduce work and information overload. *Commun. ACM*, 37(7):30–40, 1994.

[4] Alexandros Nanopoulos, Dimitrios Katsaros, and Yannis Manolopoulos. A data mining algorithm for generalized web prefetching. *IEEE Trans. on Knowl. and Data Eng.*, 15(5):1155–1169, 2003.

[5] Jean Oh, Felipe Meneguzzi, and Katia P. Sycara. ANTIPA: an architecture for intelligent information assistance. In *Proceedings of the 19th European Conference on Artificial Intelligence*, page (to appear), 2010.

[6] R. S. Sutton and A. G. Barto. *Reinforcement learning: an introduction*. MIT Press, Cambridge, MA, 1998.

[7] Neil Yorke-Smith, Shahin Saadati, Karen L. Myers, and David N. Morley. Like an intuitive and courteous butler: a proactive personal agent for task management. In *Proc. 8th Int. Conf. on Auton. Agents and Multiagent Systems*, pages 337–344, 2009.