

The Early Stages of T&Browse: A Spoken Shared Internet Browser for the Blind

Emad Eldin Mohamed and Hesham M. Kamel
College of Information Technology
United Arab Emirates University
Al-Ain, United Arab Emirates
emohamed,hesham@uaeu.ac.ae

Abstract

This paper introduces a shared Internet talking browser for the visually impaired called Talk & Browse (T&Browse) that allows a group of blind users to access Internet content and view the same HTML documents at the same time. Within the browsing session, the system allows two types of users: a sighted person (session supervisor) and blind users (session guided users). When the supervisor browses the Internet, the web pages he/she browses are then displayed on the guided users' computer screens via voice synthesis. T&Browse can be used within classroom settings where the supervisor is the instructor who can guide the users (students) through the Internet. Also, the supervisor can use the system remotely (i.e., the supervisor and the guided users need not to be placed in the same geographic area), thus promoting distance learning. To make it platform independent, T&Browse is implemented using Java.

Keywords: visually impaired, shared Internet browser, Collaborative applications

1. Introduction

Providing blind people with access to the computing and networking facilities has been faced with hurdles in recent years. A common solution to this difficulty is to assign a mentor for every user to help him/her access the contents of these facilities. This solution may not be the optimal in many cases. For example, it is usually difficult to assign a mentor

for every blind student in large classes. Moreover, coordination among the mentors within a given class may be problematic. For situations such as the one we have just mentioned, other more practical solutions have to be explored.

In this work, we investigate the special needs of blind people from computer systems with a focus on Internet content access. We propose a solution that enables the visually impaired to access Internet content through the use of a spoken shared Internet browser called Talk & Browse (T&Browse). Our proposed solution has two main features. First, it can be considered a collaborative system, since it enables a group of users to share HTML documents under the supervision of one person. Second, using a text to speech engine [11], the system synthesizes the voice of the text extracted from the shared HTML documents.

T&Browse can be used in classrooms as a standalone application to share Internet content or it can be integrated within other collaborative systems such as distance learning and teleconferencing. Moreover, despite the fact that we focus on Internet browsing, the solution we propose can be easily adapted to many other computer systems.

Motivations of T&Browse are many. First and foremost is to guide blind computer users to comprehend Internet content. Second is to enhance collaboration between Instructor and his/her students in a classroom setting. Third motivation is to promote distance learning. Fourth is to enhance communication between blind and sighted people and open another door for minority (namely blind people) to become part of society.

The proposed system tackles two main issues. The first issue is how to share Internet content among several users at the same time. A shared web browser has to deal with issue such as floor control (who controls the Internet session) and synchronization (how to deliver the same HTML document to all users at the same time). The second issue is how to integrate a text to speech engine within the system to render HTML text and other data format that can be transformed into voice. This issue proves very tricky especially when we deal with web pages that include frames and graphics. The system in its current stage deals only with the text of the documents and leaves the handling of all other data media to future work.

The rest of this paper is organized as follows. Section 2 discusses some of the related work. Section 3 presents the description of the proposed system. Section 4 presents alternatives for system design. Issues and difficulties facing the system are the subject of Section 5. The conclusion and future work are given in Section 6.

2. Related work

The shared spoken Internet browser introduced in this paper is an example of collaborative systems. A collaborative system can be defined as a system that allows a group of users, not necessarily in the same place, to work and cooperate towards the success of a common task [5, 8]. Collaborative systems mainly fall in one of two categories: collaborative aware and collaborative naïve. Collaborative aware systems are built with sharing taken into consideration, whereas collaborative naïve are single user systems that are shared using some sharing engine. Examples of collaborative aware systems include shared whiteboard and shared editor [4, 6]. Examples of sharing engines are XTV [1], which shares single user X-Window applications, JCE [2], which shares Java applications, and Microsoft NetMeeting [9] to share MS Windows applications.

Earlier systems that share Internet content can be found in [7, 10]. In these two systems, many features such as floor control and tele-pointer have been included. These two systems, however, do not support voice synthesis.

3. System description

As shown in Figure 1, T&Browse consists of two main parts; the first is for the supervisor and the second is for the guided users. The supervisor is connected by TCP connections with all users.

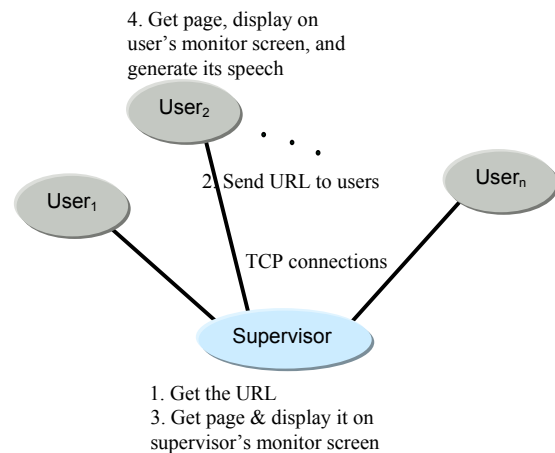


Figure 1. System description

Initially, the supervisor starts a browsing session and waits for users to connect with him/her. After connections are established, the supervisor can browse the Internet and access its contents (step 1 in the figure). In step 2, the supervisor program sends all the links the supervisor visits to users. The supervisor then downloads the required HTML document in its computer (step 3). In step 4, user programs download the HTML documents of these links on their computers. For every downloaded HTML documents, the user program parses the document to extract the text from it. The system then passes the extracted text to a text to speech (TTS) engine that synthesizes the voice of the given text. All users hear the speech of the document provided by the supervisor at the same time. In this sense, the system provides access to the same information on the Internet delivered with voice synthesis to several users under the guidance of a supervisor.

T&Browse in its current stage is implemented as a Java application to make it platform independent. It makes use of a Java text to speech engine called FreeTTS, which is based on a text to speech engine developed at Carnegie Mellon University [11].

4. Design alternatives

The system in its current implementation uses a client/server connection between the supervisor and the users to send control information from the supervisor to the users. These control information mainly includes the URL of the pages the supervisor is visiting.

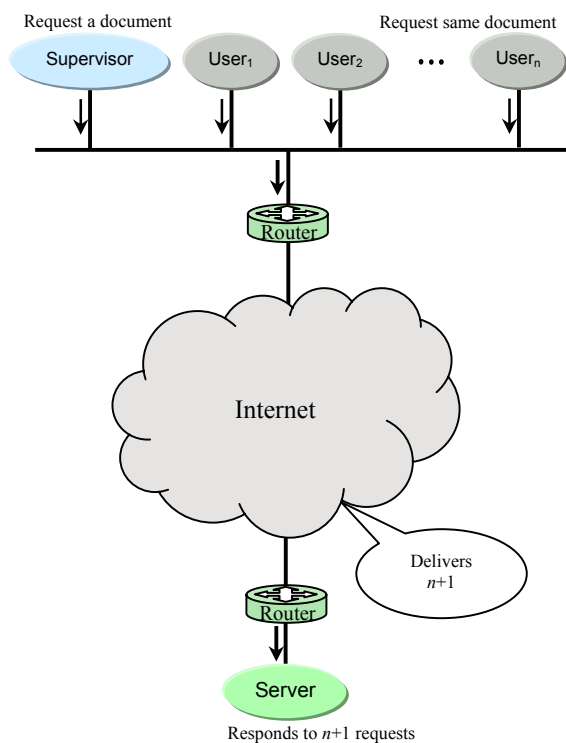


Figure 2. Multiple unicasts

The system requires the users to directly download the HTML documents from their URL servers. Figure 2 demonstrates how this technique works. As the figure shows, every user downloads the same HTML documents. If we have n users, the network routers as well as the server hosting the document respond to $n + 1$ requests. In addition to overloading the network routers and the server this technique may create a synchronization problem; some users may get the HTML document earlier than others.

Some other design alternatives exist. One alternative is to let the supervisor download the HTML document and unicast it to every user. This solution works well if all users exist in the same LAN. At the positive side, this technique optimizes the use of the network and the server of the HTML document. At the negative side, it overloads the computer of the supervisor and its network connection to all users making the supervisor load of the order of $n + 1$, where n is the number of users.

A solution to the above problem is to use multicast communications. In this case, the supervisor multicasts the HTML documents to all users. Since the HTML document must be communicated reliably, a reliable multicast layer is required. The current implementation of IP only supports best effort multicast. Many third

party reliable multicasts exist. Their overhead on the sender and receivers, however, approaches that found in unicast communication.

One more design alternative is to let the supervisor download the HTML document, extract the text from it, and communicate the extracted text to the users. Upon the receipt of the text, a user performs voice synthesis for the received text. This way, users may not be able to display the HTML document appropriately on their monitor screens as formatting information along with all graphics are not sent. This approach, however, serves two purposes. First, it improves network usage, as the size of text normally is much less than that of the HTML documents. Second, it improves synchronization among users, as all users get the text almost at the same time and they are left only with voice synthesis.

In addition to extracting the text from the HTML documents, the supervisor can also perform voice synthesis and send the generated audio to users. This may load the supervisor computer and increase network traffic since the size of audio data may well exceed that of the text. However, this technique has an important advantage: it supports a light weight user program as most of the computations are moved to the supervisor. This in effect allows us to implement the user program as an applet that can be loaded easily from the Internet. Each applet establishes a connection with the supervisor and receives the speech of the HTML documents visited by the supervisor.

5. Issues and difficulties

Implementing T&Browse, we have been faced with several issues and difficulties. The main three issues, however, are synchronization, floor control, and text to speech integration. The following is a discussion of these issues.

5.1. Synchronization

An important factor that affects the overall success of the proposed system is synchronization. The system is said to have good synchronization if the variance of the transmission delays experienced by the users is within a reasonable limit; that is users receive data at around the same time. Having a good synchronization results in a better interaction among the users and the supervisor.

Synchronization can be achieved in several ways. First the design alternative appropriate to the case in hand should be selected. For example, if the supervisor and all users exist in the same LAN (in a classroom,

for example), then it is better to require the supervisor to multicast the HTML document to the users. If the users are located sparsely in a large network (as in typical distance learning systems) it may be better to let every user download the HTML document from its original source. In this case, synchronization can be improved by measuring the longest delay the users experience and delay other users accordingly; a difficult task if the number of users is large.

5.2. Floor Control

The second issue to consider is floor control. Floor control decides who has the control over the browsing session at a specific time [3]. The current system implementation assumes control only to the supervisor. However, enhancements to system can allow the users to control the browsing session floor. In this case, a floor control manager must be included to resolve simultaneous floor control requests.

There are two main approaches to manage floor control: centralized and distributed. In the centralized solution, a central server receives requests to control the floor and grants the requested control according to some policy the server has. An example of a distributed solution is to have a token that circulates around the supervisor and users in a predetermined order. The user keeps the token if it desires to control the floor; otherwise it passes to the next user (supervisor).

5.3. Text to speech integration

Given the HTML document, a user has to synthesize voice from the text of the document. The first step is to parse the document and extract the text from it. The second step is to generate speech from the extracted text. The current implementation of the system uses FreeTTS, a text to speech engine based on a system developed at Carnegie Mellon University [11]. For graphics rich document, the extracted text may not help much in expressing the content of the document. This poses a challenge as it renders the system useless for this kind of pages.

6. Conclusion and future work

This paper has introduced T&Browse; a shared spoken Internet browser. T&Browse allows supervisor to guide a group of users through the Internet so that they can access its contents. System architecture along with other design alternatives have been presented. Main issues and difficulties faced when developing the

system have been discussed. Of particular interest is how to represent graphics rich HTML documents to a group of blind users. This issue is left for future work.

References

- [1] H. Abdel-Wahab and M. Feit, "XTV: A Framework for Sharing X Window Clients in Remote Synchronous Collaboration," Proceedings, IEEE TriComm '91: Communications for Distributed Applications Systems Chapel Hill, North Carolina, pp. 159-167, April 1991.
- [2] H. Abdel-Wahab, B. Kvande, O. Kim and J.P. Favreau, "An Internet Collaborative Environment for Sharing Java Applications," Proceedings of the 5th IEEE workshop on Future Trends of Distributed Computing Systems, Tunis, Tunisia, pp. 112-117, October, 1997.
- [3] H.-P. Dommel and J.J. Garcia-Luna-Aceves, "Floor Control for Multimedia Conferencing and Collaboration," *Multimedia Systems Journal (ACM/Springer)*, Vol. 5, No. 1, pp. 23-38, January 1997.
- [4] C. Ellis, S. Gibbs, and G. Rein, "Design and Use of a Group Editor," Proceedings of the IFIPTC2/WG2.7 Working Conference on Engineering for Human-Computer Interaction, Napa-Valley, CA, USA, 21-25 August 1989, North Holland, Amsterdam, pp. 13-28, 1990.
- [5] C. Ellis, S. Gibbs, and G. Rein, "Groupware: Some Issues and Experiences," *Communication of the ACM*, Vol. 34, No. 1, pp. 38-58, 1991.
- [6] R.S. Fish, R.E. Kraut, M.D.P. Leland, and M. Cohen, "Quilt: A Collaborative Tool for Cooperative Writing," In R.B. Allen (ed.) *Conference on Office Information Systems*, March 23-25, 1988, Palo Alto, CA, USA, vol. 9 no. 2, *SIGOIS Bulletin*, ACM Press, New York, pp. 30-37, 1988.
- [7] S. Greenberg and M. Roseman, "GroupWeb: A Groupware Web Browser," Proceedings of Video ACM CSCW'96 Conference on Computer Supported Cooperative Work. Boston, USA, ACM Press, 1996.
- [8] I. Greif (ed.), "Computer-supported cooperative work: a book of readings," Morgan Kaufmann, 1988.
- [9] Microsoft NetMeeting,

URL://<http://www.microsoft.com/netmeeting/>

[10] E. Mohamed, " SBrowse: A Collaborative Shared Internet Browser," The Sixth Annual U.A.E. University Research Conference, Al-Ain, United Arab Emirates, 2005.

[11] W. Walker, P. Lamere, and P. Kwok, FreeTTS: A speech synthesizer, URL: <http://freetts.sourceforge.net>