On-Line Courses Come Alive

Ranjana Shukla
UNITEC Institute of Technology

Hira Sathu
UNITEC Institute of Technology

Zhong Tang
UNITEC Institute of Technology

Follow this and additional works at: http://aisel.aisnet.org/amcis2001

Recommended Citation
http://aisel.aisnet.org/amcis2001/226
ON-LINE COURSES COME ALIVE

Ranjana Shukla  
UNITEC Institute of Technology  
rshukla@unitec.ac.nz

Hira Sathu  
UNITEC Institute of Technology  
hsathu@unitec.ac.nz

Zhong Tang  
UNITEC Institute of Technology  
ztang@unitec.ac.nz

Abstract

The objective of the paper is to investigate techniques to provide a choice of multidimensional interaction possibilities between student-student and student-teacher at an academic institute. The technologies adopted take into consideration the constraints posed by Internet infrastructure that is used as the ubiquitous means for this interaction. The research adopts a phased approach to study and overcome the constraints posed by the existing communication infrastructure. In Phase one, current tools and software packages were studied with a view to integrate voice and image functionality into the Course Info Server. The Phase 1 results presented in the paper indicated the possibility of using voice as an alternative to text chat. However, as for the quality of voice chat the results have not been encouraging since the voice deteriorates when the text chat is sent simultaneously. The concept and the approach for Phase 2 is included in this paper. In this phase it is planned to transmit only text, which is a standard Course Info Server feature. At the receiver's end provision for reception of text and audio both are envisaged. This is achieved by conversion of the text to audio at the receiver end and provides an electronic voice to accompany the usual text message. In the third phase it is intended to integrate voice recognition software to convert voice into text and then send it over the Internet. At the receiver’s end a reverse process of conversion from text to speech as done in Phase 2 is adopted. The final phase intends to use the methodology of speech to text and text to speech conversion techniques and wed these to near natural voices of the speakers locally, without travelling over the Internet (voice patterns obtained from the local CD).

Introduction

The lack of personal touch or the humane approach in online or web-based learning is a major impediment to its growth. Keeping in view the other advantages those accrue, the only option is to integrate features that would bring in the human touch to the learning process. Integration of a variety of communication media is expected to make the on-line courses more friendly, interactive and interesting. The other synergistic benefits like a feeling of togetherness, that you are not alone in your pursuit of knowledge, and there are others who share similar interests and the like would also accrue.

The Problem

In this research the authors attempt to research the additional functionality that could be added to the Black Board Course Info Server used at UNITEC Institute of Technology, Auckland. Students use the services of this Course Info server for emails, course / group discussions, and some on-line chat for learning outside of classroom contact hours. It is well established that textual data when complemented by voice and video would bring alive an on-line learning process. Without these features this communication very often is a dull learning process. However, the addition of voice and video to text data requires a large bandwidth as well as guaranteed quality of service. It is well established that use of Internet is ubiquitous for on-line learning by the academic community at UNITEC, as elsewhere. The availability of these guarantees of service over the Internet is still a far cry. For experienced users of various Internet Phone applications the unreliability of the voice quality is not unusual. The glitch noticed during news reading on the net is a case in point. As pointed by Minoli and Minoli (1998) a possible solution could be better management of corporate and IP networks to handle real-time network traffic needs. This entails initiatives by way of prioritisation, wider bandwidth channels with adoption of feature rich protocols like ATM or near ATM capabilities. These initiatives are not only expensive but beyond the control of the user communities involved. Quoting Minoli and Minoli (1998)
"Up to now, vendors have been of the small upstart category, rather than mainstay…Of late, high tech leaders such as Cisco, Lucent and Nortel have shown interest in entering the market”. Therefore, other alternatives need to be investigated.

The Approach

This research adopts a phased approach with specific goals at the end of each phase. There are four phases for the research and implementation of this research. Brief details of these in order of their development and integration into the Blackboard Course Info Server are covered here.

Phase 1

This part of the research is conducted with groups of students who use Blackboard Course Info Server for their group assignment. For this study students are divided into groups of four to five students each. The whole group goes on line at a pre-specified time of the day every week for about an hour. The on-line session starts with a presentation from one of the designated students followed by a question-answer session. During the on line discussion photographic images of every student in the group are up on everyone’s screen. The image of the person who is on line is highlighted and is complemented with the person’s voice. Any student asking or answering a question is also broadcast (multicast-group) over the Internet in a similar fashion. In this phase the authors undertake the study of current tools, and software packages with a view to integrating voice functionality into the Course Info Server. This research is expected to cover NetMeeting and Paltalk, at the minimum and other related software later if possible.

Review of Voice over IP Applications

The benefit of voice over IP applications for distance learning is well appreciated (Minoli 1996). As part of Phase 1 review of available software packages in this area was carried out. These included MSN Messenger, Microsoft’s NetMeeting, PalTalk, Dial Pad, MediaRing, and VocalTec Inc. Internet Phone.

Most of the software packages are quite general in nature and are aiming at the broader audience. Some of them have instant messaging features and there are others that provide free or nearly free net telephone facilities basically between two parties. As per a VocalTec press release (November 12, 1996), its Internet Phone allows up to 100 people to participate in an audio conference when used in conjunction with the VocalTec Conference Server. Lucent Technologies has been a leader in server-based approach to Voice over IP. It also enables audio and video conferencing over the Internet (Lucent- web site) but is positioned primarily for the carriers. Better ones now come with video capability too, an extremely powerful feature from the human angle. Some software empowers the user with multipoint ‘chat’ option, file transfer, file share and whiteboard capabilities.

The telephone conversation over the net is not of a very good quality but since it is free or nearly free it is used by large number of Internet users. According to Yahoo! (Yahoo! 1997), VocalTec Telephony Gateway (PC to Phone) dominated the consumer and carrier segments of the Internet Telephony market in 1996, but of late the market is shifting in favour of other technology providers. The initial study helped to narrow down to two products that had similar features to what the project team were looking for, a live class room (one to many dialogue capability, besides other available functions). Very brief features of NetMeeting and Paltalk are described here.

NetMeeting

NetMeeting, developed by Microsoft, is very good software with good documentation. One can download the NetMeeting developer’s kit, which is a collection of sample code modules. A developer can modify and integrate the appropriate models in his/her own software product. It has the following features: placing calls on multiple users with an ability to organise own user
groups, receiving a call when one is running a net meeting session and is contacted; sharing of program with all other members of the group. Where program control is given to other members of the group, other members of the group as well can see and modify the program. Everyone participating in the group can simultaneously draw and see on the whiteboard what others have drawn. As for Text Chat, all members of the chat group can be addressed. Voice functionality with sound quality depending upon the sound card quality, CPU speed and other factors. A facility exists to use Chat and Whiteboard with many but the voice call is restricted to only one person at a time. The voice chat is full duplex (like ordinary phone conversation). The video feature, like audio, is applicable between two users at a time.

PalTalk
Software can be downloaded from the PalTalk.com site for free. Most of its features are similar to NetMeeting. Major features that interested the authors were its audio and video conferencing feature suited for group interaction. Its ease of installation was also a bonus. This was finally selected for use.

Interface development
The groups of students undertaking the assignment in Voice over IP from the Data Communications and Networks-2 (DCN-2) class were expected to study the feature of audio conferencing over the Internet among other things. To maintain consistency in the environment for study and the observation thereof, the authors developed an interface to the Blackboard Course Info Server for audio and text conferencing with images. Use of “PalTalk” as considered above was made for audio and chat conferencing. The main features of the interface developed are discussed briefly below.

- **Help:** This feature basically describes how students can use the audio, text conferencing with imaging feature from the Blackboard Course Info Server efficiently.
- **Loading and Installing PalTalk:** The centralized loading and a common source of instructions for installing and configuring are with a view to maintaining a common and a consistent environment for the study across the student groups.
- **Group and member classification:** This covers the classification of students from the DCN-2 class into groups of 3 to 4 students undertaking the Voice over IP assignment project.
- **Group Authentication:** This feature is incorporated to ensure the privacy of the audio and text chat conferencing of each group. The students using audio conferencing were given group password to enable this.
- **Database:** A database was created to keep the student information. It has group number, student name, and group passwords, group leader and student’s image fields. This database helps in easy accessing and updating of student data.
- **Images:** The interface program also brings up images of all students in a specific group on to the desktops of individual students.
- **Highlighting group leader’s image:** Group leader’s image is highlighted to facilitate the session.
- **Starting PalTalk:** PalTalk is invoked from the interface but is resident on the individual’s machine.

**Phase 2**

The authors, having experienced in Phase 1 the results of integration of voice and its broadcast over the Internet along with the text chat, would now be able to quantify the often-expressed limitation of sending voice over IP. Therefore, in this phase it is planned to transmit only text, which is a standard Course Info Server feature. However, at the receiver’s end provision for reception of text and audio both are envisaged. This is achieved by conversion of the text to audio at the receiver end and provides an electronic voice to accompany the usual text message. Thereby, it will be possible for the receiver to hear the message too, though only a replication of the sender’s real voice. An interface will be developed for incorporating text to speech conversion for adding of the audio component to the text of the Blackboard Course Info server. This approach will not have any bandwidth problems and latency issues as faced in voice transmission over IP, as the voice is enabled (text to speech conversion) only at the receiver’s end and not transmitted over the media. Implementing this will involve experimenting with text to speech conversion software packages and integration of the selected software into the participant machines.

**Phase 3**

In this phase of the research a two-way speech capability is trialled. The authors will integrate speech to text conversion software at the source end to enable a student/user to speak instead of keying in text. This speech would then go through speech to text conversion software. Thus the voice is converted into text and then sent over the Internet. It is envisaged that some amount of text correction may be required before transmission. This is on account of the present speech recognition technology not permitting...
100% correct conversion. Previous studies (Oviatt, 2000) demonstrate error rates increasing from 29% for a carefully read
dictation to 53% for natural spontaneous interactive speech. At the receiver’s end a reverse process of conversion from text to
speech as done in Phase 2 is taken up to complete the communication process. This phase entails research into various software
packages available for speech to text conversion, followed by the selection of the best one suited for the purpose. The next step
will be developing interface software (speech-to-text conversion) to piggyback on Blackboard Course Info Server. This is
expected to be a very crucial piece of software, since an inappropriate integration may render the entire Course Info Server non-
functional. The text-to-speech and speech-to-text software packages will reside on participant machines and will be invoked by
the interface software.

Phase 4

In the final phase it is planned to use the methodology of speech to text and text to speech conversion techniques and wed these
to the natural individual voices. Voice patterns for all the students involved in an online course of a discussion group or chat-
session, with mutual agreement, will be saved on a CD. During the discussion session whenever a text message is received from
a particular student, his/her voice pattern from the CD will be activated. The student receiving the message will now be able to
hear the message from the sender in the sender’s own voice, not an electronic voice, bringing in the much-required personal touch.
The choice of having the text message accompany the voice is left to the control of the recipient.

A lot of work is envisaged to generate student voices from respective voice patterns. This is expected to involve both technical
and ethical issues. The ethical issues are on account of the approval from the students agreeing to have their voice reproduced
from earlier samples recorded and stored. Several practice sessions may be required to perfect speech reproductions from text
related to individual speech patterns. The modalities for exact implementation are yet to be worked out. This area has great
research potential and the outcomes may find wide usage despite the additional logistics of the availability and exchange of speech
patterns.

Conclusion

The culmination of Phase 1 revealed that students find voice and image helpful in the communication process. However, the
existing half-duplex voice was poor with regard to interaction. The simultaneous use of text chat along with voice faded the voice
quality. The availability of voice as an option was appreciated in the feedback especially with the higher age group tutors with
poor keyboarding skills. The addition of voice and image components was considered to improve the overall communication
process between the students as compared to the existing plain text. The metrics for this have been the student and the tutor
feedback questionnaires.

The results of the first phase have motivated the authors to attempt Phase 2 with the underlying methodology: build and operate
a chat server over which the research team has complete control unlike PalTalk used in Phase 1. This web server will have at the
minimum, Exchange Server Utilities like MS Chat Service and Internet Locator Service installed. The web interface developed
for Phase 1 will include an additional module, developed using MS NetMeeting S.D.K. or a comparable product to enable text
chat. Once the text chat between the groups is tested and proved, the client group members will install Microsoft Text-To-Speech
engine (TTS). The final part of this phase would be to integrate the above interface with TTS engine and enable text-to-voice
conversion. The developed interface will also need changes to the database developed earlier for Phase 1, to include the gender
and speech characteristics like pitch and speed. It is proposed to use similar metrics for Phase 2 as used in the first phase.

References

Oviatt, S., “Taming Recognition Errors with a Multimodal Interface,” Communications of the ACM, September 2000/Vol. 43
No. 9, pp445-51