

MSc projects 2008

Barry Cheetham

BC08MSc.1: Estimating network QoS parameters & simulating video conferencing links.

This project is concerned with interactive real-time audio/video conferencing as may be applied to distance learning over the Internet.

The first phase of the project is an investigation of the "quality of service" (QoS) parameters which quantify the performance of a real Internet link for conferencing, and the accurate measurement of these parameters. A means of measuring QoS parameters such as round-trip delay, jitter and packet loss probability should be devised and used to obtain measurements for a number of different Internet connections, geographical locations, times of day and packet sizes.

The second phase of the project will design and implement a simulator that emulates different QoS behaviour based on the measurements obtained during the first phase of the project. This will allow distance learning systems which involve interactive real time communication by Internet to be tested "in house" over typical Internet links before the systems are released to world-wide participants of distance learning courses.

BC08MSc.2: Damaged packet delivery and PLC for VoWLAN

This project concerns the use of a wireless LAN to provide interactive voice communications, i.e. 'voice over WLAN' (VoWLAN), as well as data. 'Voice over IP' (VoIP) is now widely used for telephony over wired networks and is being increasingly used over wireless networks despite the fact that wired and wireless computer networks were designed primarily for data. Measures which aim to provide or guarantee a certain 'quality of service' (QoS) to telephone traffic are often introduced into 'medium access control' (MAC) mechanisms. The presence of bit-errors in a received packet may be detected by a 'cyclic redundancy check' (CRC) and, in wireless networks, forward error control (FEC) is used to allow the receiver to correct a small number of bit-errors by means of a 'Viterbi' FEC decoder. Despite these measures, voice packets may be irreparably damaged, lost in transmission or excessively delayed so that they are too late to be useful. 'Packet loss concealment' (PLC) is used to replace lost or excessively delayed voice packets by artificially created voice packets in such a way that the distortion, as perceived by a listener, will be as small as possible.

The QoS provided by wireless networks will normally be different from that provided by wired networks in that irreparable damage will be more common. Excessive delay will be less common with wireless. With interactive telephony, it is normally not possible to use transport layer acknowledgements and retransmissions when errors occur as with TCP. Currently, with VoIP and VoWLAN, irreparably damaged packets are discarded and replaced by PLC created packets, even when the number of uncorrected bit-errors is small.

The aim of this project is to investigate how packets that are irreparably damaged by a wireless LAN can be made use of if they are not discarded. Packets are considered

irreparable when the FEC Viterbi decoder fails to correct all the bit-errors that occurred during transmission. The following questions may be asked:

- (1) What is the effect of the Viterbi decoder on irreparable packets?
- (2) If uncorrected bit-errors are simply ignored, under what circumstances is the distortion in the voice sound worse than that created by throwing the whole packet away and using PLC to produce a substitution?
- (3) Since WLAN cards can perform automatic retransmissions at the DLL layer and there may be several irreparably damaged versions of the same packet available at the receiver, how can these versions be used to produce a packet with fewer or no bit-errors ?
- (4) If irreparably damaged packets are to be made use of rather than simply discarded, is the 'convolutional coding' and Viterbi decoding mechanism used currently by wireless LANs the best option for voice?

The investigation may be carried out by means of software simulations using MATLAB or another language.

BC08MSc.3 : Viterbi decoder for forward error correction in mobile networks.

There are two possible projects concerned with correcting bit-errors in mobile LAN equipment. Both involve a study of the Viterbi decoding algorithm for convolutional coding.

The aim of the first project is to investigate a power saving strategy whereby the decoder is de-activated when bit-errors are not occurring. A new way of doing this has been devised but remains to be fully investigated.

The aim of the second project (probably more challenging) is to study 'forward error correction' (FEC) as used in IEEE802.11 wireless LAN technology and elsewhere. The IEEE802.11 standard employs 'convolutional coding', a cyclic redundancy check and retransmissions to accommodate bit-errors that will usually occur on wireless links. The convolutional decoding is achieved by a 'soft decision input' Viterbi decoder, and any packets that fail the cyclic redundancy check after the best efforts of error correction are normally discarded. The aim of this project is to understand the principles of a 'Viterbi' convolutional decoder which has 'soft decision' outputs as well as soft decision inputs, and to apply this to improve the FEC performance of a wireless LAN. A well known algorithm is known as 'SOVA', but it is intended that we develop our own version, perhaps in MATLAB. This project is related to BC07MSc2, but will concentrate more on the algorithm than the VoIP application.

Further detail: The contents of WLAN packets are convolutionally coded to allow a degree of bit-error correction if errors occur. The convolutional decoding at the receiver is efficiently done by a Viterbi decoder which these days is usually a 'soft decision' input Viterbi decoder. 'Soft decision' input means that the receiver will assign a level of confidence to each bit it receives: '111' could represent 'definitely 1'; '000' could represent definitely '0', with '110', '101' and '100' representing 'probably 1', 'perhaps 1' and 'guess 1' respectively. Similarly for '001' etc. A receiver expecting say 1 volt or 0 volt pulses would make these soft decisions according to how close the noise affected received voltage is to 1 volt or zero volts. A conventional Viterbi convolutional decoder will output a 'hard decision' bit stream

which is hopefully correct. The bit-stream will likely include a 'cyclic redundancy check' (CRC) allowing the receiver to have a good idea that the data is wrong when it is wrong. A conventional WLAN receiver will just discard Viterbi decoded data when it fails its CRC check. If we are to consider making some use of the data in packets which fail the CRC, i.e. damaged packets, rather than just discarding them the use of a Viterbi decoder whose output as well as the input consists of 'soft decision' bits, may be beneficial. The aim of this project is to study convolutional coding, soft decision decoding and the many available decoding algorithms. Ultimately the requirement is a demonstration of a working soft decision output Viterbi decoder and an assessment of its effectiveness for the purpose describe above. This project is strongly related to a previous one but is focussed specifically on the soft output decoder.

BC07MSc4: Effect of bit-errors on compressed speech and music.

Speech, music and multi-media are generally compressed for transmission over wired and wireless networks to reduce the required bit-rate. The effect of bit-errors tends to become more serious the more the compression. The aim of this project is to investigate the effect of bit errors as may occur in Mobile Communication over a WLAN, on the quality of the sound when a number of different compression techniques are used. The G711 (64kb/s), G726 (32kb/s) and G729 (8 kb/s) standard compression techniques for 'narrow band' speech should be investigated along with MP3 for music. The project could be extended to pictures. The aim is to produce an application for demonstrating the effect of bit errors and the effect of commonly used forward error correcting (FEC) coding. The use of software to automatically evaluate 'mean opinion scores' (MOS) for narrowband speech will be part of this project. The MOS software and the speech and music coding techniques will be available, but something of their underlying principles must be understood.

BC08MSc5: Speaking or singing voice trainer

These are two possible projects concerned with the human voice as used for speaking and/or singing. A study of the mechanisms by which humans speak, sing, convey emotion and generally communicate reveals an anatomical process that, in outline, is fairly simple to understand and can be modelled by a computer program. An understanding of sound perception and the mechanism of the ear allows the computer modelling of voice to be made even more effective. Mobile phones use such a model to produce the speech you hear and this allows speech to be transmitted at low bit-rates to preserve radio spectrum. However, an understanding of the human voice can have much wider applications than bit-rate compression.

Mastering some of the basic principles of human speech and its perception is the goal of the first part of each of these projects. Attending Comp30291 'Digital Media Processing' in the first semester and learning MATLAB would be a definite advantage.

The two possible projects are as follows:

'Speaking voice trainer'. Voice training as used by actors and politicians to enhance their presence and persuasiveness may be studied and a computerised voice analysis package may be used to help this process. A famous example is the training Mrs Margaret Thatcher, British Prime Minister from 1979 to 1990, received to lower the pitch of her voice to give herself greater 'gravitas' and authority. Voice experts teach that the voice must 'resonate' deep within the body (as with Winston Churchill) rather than within the mouth (David Beckham) or the nose (Michael Cain). Measuring the pitch of the voice by computer is an interesting and achievable goal, and the detection of body resonance may be possible by examining some Fourier spectral graphs.

The inspiration for the speaking voice trainer came from a comedy program on BBC7 called "I'd know that voice anywhere" presented by Alister McGovern (First broadcast 19 Jan 2000 but repeated many times). As well as raising the points mentioned above, it discusses ventriloquism and features Rory Bremner and others who can mimic other voices. 'Voice morphing' and the creation of cartoon voices was not mentioned but is related. Anyone interested in these aspects of voice analysis and/or synthesis could adapt the 'speaking voice trainer' in this direction. Downloading the BBC7 program would be a good start (and a good laugh).

The main aim of the 'speaking voice trainer' project is to develop an application that provides feedback to a person wishing to train his or her speaking voice. Specific objectives are as follows:

- (i) To consult professional publications on speaking voice training and, if possible, to contact staff in the University or elsewhere.
- (ii) To study the physical mechanisms responsible for voice production.
- (iii) To develop tools for analysing sections of the voice to determine pitch, intonation and other characteristics.
- (iv) To test the analysis tools by applying them initially to famous recordings, for example of Margaret Thatcher, Winston Churchill and Michael Cain.
- (v) To develop the tools into a prototype 'speaking voice trainer' providing a suitable 'graphical user interface' to the user and convenient feedback from the analysis.
- (vi) A suggestion is that the user be asked to repeat a phrase or sentence in a particular style and be informed how close he/she is to the style.
- (vii) The usefulness of the application must be evaluated with some subjective testing.

The willingness to come to grips with some basic digital signal processing techniques and to do some high level language programming with GUIs is needed.

'Singing voice trainer'. This project has much in common with the speaking voice trainer and would be ideal for somebody who sings or has sung in a choir. The aim is to develop and test a computerised 'singing voice trainer' and to try it out on a real singer to determine whether it is of any use at all. The objectives are as follows:

- (i) To develop software that checks that a singer is singing the right notes at the right time. It could have a copy of the score (with words), a few bars would be sufficient to begin with, and make a recording of a person's

- attempt to sing the words to the notes. Segmenting the person's recording into individual notes presents a challenge. Then the program could 'Fourier analyse' the frequencies sung and check that they are correct.
- (ii) The timing may now be carefully analysed, checking for late or early entries and the singing of quavers as crotchets, for example (a common mistake).
 - (iii) The software could then go further, for example by looking out for the dreaded early or late consonant. Since consonants like 's' and 'sh' are unvoiced and are therefore not periodic (the waveform is more random or noise-like) they are unable to carry a pitch. Vowels carry pitch as they are periodic. Mistimed consonants (especially 's's) are an anathema to conscientious choral directors and singers must practise to get this right.
 - (iv) Resonance and projection are as important with singing as with speaking, if not more so, and an analysis of these aspects of voice recordings could be made.
 - (v) Professional singers maximise sound volume by modifying vowels to make vocal tract resonances coincide with pitch harmonics and it would be fascinating to be able to analyse this process and maybe help a singer to emulate it.
 - (vi) The prototype computerised 'singing voice trainer' should be provided with a 'graphical user interface' to allow it to be tried it out on a real singer to determine whether it has any potential.

BC08MSc6: An application for teaching Westerners to recognize and reproduce tones in Chinese

For many Westerners studying Chinese, the recognition and reproduction of the four basic tones in Mandarin, and the many more in Cantonese, are a source of great difficulty. Some 'wai-guo-ren' simply cannot discern the differences as they are so attuned to other aspects of intonation. Reproducing the tones can be an even more difficult problem. Imagine the frustration of a poor old Englishman who has learnt his vocabulary only to find that no Chinese person has a clue what he is saying because the tones are all wrong.

This project is for a Computer Scientist with knowledge of Chinese, preferably Mandarin, and an interest in developing interactive educational software which includes sound analysis and reproduction.

Mastering some of the basic principles of human speech and its perception will be required and attending Comp30291 'Digital Media Processing' in the first semester and learning MATLAB would be an advantage.

The main aim of the project is to develop an application that demonstrates the use of Mandarin tones to a Westerner, analyses his/her attempts to reproduce Chinese words and sentences correctly, and provides feedback and suggestions for corrective action. Specific objectives are as follows:

- (i) Consult professional publications on the teaching of Chinese to foreigners and, if possible, to contact language teaching staff in the University or elsewhere to gain from their experience. The teaching methods adopted by Chinese teachers of foreigners and the computer software packages currently available should be considered and references on automatic speech recognition as applied to tonal languages should be researched as background.
- (ii) Study the physical mechanisms responsible for voice production.
- (iii) Develop tools for segmenting recorded sections of spoken Mandarin to isolate single words. Then devise the means of measuring the fundamental frequency of the voice at suitable intervals to determine how the pitch of the voice varies within each word.
- (iv) Given a 'pin-yin' representation of the segment of spoken Chinese, the application should then be enabled to impose the standard symbols over the vowels to indicate the tones (right or wrong) as detected by the application. These can be compared with the correct tones and imperfections in either the software or the speaker can be identified.
- (v) Analyse the tones of a clear native speaker of Chinese to verify the software before moving on to a trainee.
- (vi) Build the tools into an application which produces written Chinese (probably in pin-yin), invites the user to speak the words or phrase, and then analyses the result. The correct pronunciation could then be given as sound output.
- (vii) Introduce the option of producing Chinese words and inviting the user to identify the tones via a keyboard input.
- (viii) Develop a suitable 'graphical user interface' with convenient feedback from the analysis.
- (ix) Subjectively test the software on a user and, by seeking the advice of a professional teacher, give an assessment of its potential for further development.

DO NOT USE THIS ONE IN 2008:-

BC08MSc.7 : Investigating voice over wireless LAN protocols

The aim is to simulate a cordless voice telephone network which uses a wireless LAN (IEEE802.11) scheme rather than a conventional GSM cellular system. This allows a private telephone network (PABX) as may exist in a small commercial company to operate alongside and interact with data terminals such as computers and PDAs. Various 'voice-over-IP' schemes have been proposed in the research literature and these should be evaluated to discover roughly how many speech users may be accommodated on a single wireless LAN. It should then be possible to propose more efficient 'voice-over wireless LAN' (VoWLAN) schemes without the normal overheads of IP to accommodate yet more speech users with better speech quality. Many voice over wireless LAN techniques propose to use the point co-ordination function (PCF) of IEEE802.11 though Liu and Wu propose a novel distributed co-

ordination function (DCF) mode approach which uses the IEEE "beacons" and power saving facility (MAC approach) to achieve a "pseudo-time-division-multiplexing" technique for each speech channel. Many other researchers conclude that PCF would poorly support voice and are now looking at the problem of integrating voice with 'contention mode' wireless networks and exploring the new MAC sub-layer of the proposed IEEE802.11e standard. Popular MAC sub-layer approaches include "distributed fair scheduling" (DFS) and "blackburst".

NB This project will probably be restricted to VoIP and the use of standard IEEE802.11 protocols in the normal 'contention mode'.