Mobile Systems

Lecture 9

Voice & multimedia communication & more on CDMA

COMP28512

Steve Furber & Barry Cheetham

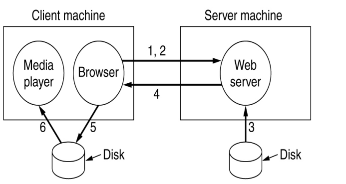
14 Apr 2015

Multimedia communications

- Needs to move large amounts of data
  - efficiently, reliably, sometimes in real-time & with delay limit
- Compression
  - Used to reduce data volume
  - Allows quality/volume trade-offs
- Bit-error correction
  - Used to improve reliability
  - Increases data volume by introducing redundancy.
  - Allows reduced transmit power
  - Must be adapted to channel characteristics
- Power optimization
  - There are many trade-offs across all aspects of the system

Mobile media: Local copy model

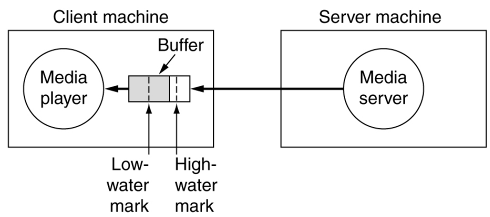
- Download file
- Use resend to recover from errors
- When complete file is available, play it.
- This is non real time – simple, but may not be ideal??



1. Establish TCP connection
2. Send HTTP GET request
3. Server gets file from disk
4. File sent back
5. Browser writes file to disk
6. Media player fetches file block by block and plays it

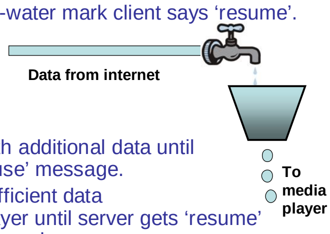
Streaming media: buffering

- Start media player as soon as sufficient 'reservoir' of data is available.
- Reduces delay & memory requirement
- Buffer smoothes out network jitter.



Buffer management

- Media server starts sending data.
- When buffer fills to high-water mark client says 'pause'
- When it empties to low-water mark client says 'resume'.
- Leaky bucket analogy!



- Any problems?
  - –buffer must cope with additional data until server gets the 'pause' message.
  - –buffer must have sufficient data to supply media player until server gets 'resume' message & new data arrives.
- Buffer overflow or underflow must be avoided if possible.

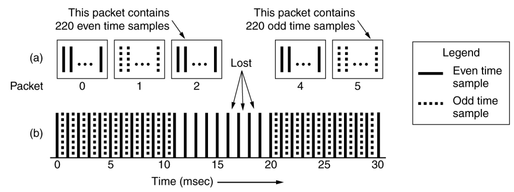
Bit-errors & lost packets

- TCP not generally used for streaming multimedia
  - Resending lost or damaged packets too expensive.
  - Effect of bit-errors less critical than for text.
- New strategies needed to cope with bit-errors
- Next slide gives a strategy for streaming voice & images.
- Becomes a little more tricky when they are compressed.

MANCHESTER 1824
The University of Manchester

Streaming media: lost packets

- Alternate samples in different packets?
  - interpolate for lost packet
  - clearly useful for uncompressed data



- **Assume we receive packets 0,1,2 but lose packet 3.**
- **Interpolate between samples of packet 2 to get approximatn to packet 3.**

Lecture 9 COMP28512 7

MANCHESTER 1824
The University of Manchester

VoIP telephony

- Different from streaming as it is interactive.
- 'Round trip' delay must be limited to ≈ 0.5 seconds.
- Otherwise people start interrupting each other.
- Cannot use TCP: any retransmissions will be too late
- Modified form of UDP used: called RTP.
- "Fire & forget" – if it's lost, it's lost!!
- Previous strategy would not work well for VoIP telephony.
  - It incurs too much delay.
  - Bit-errors can be tolerated – packet loss concealment
- Two forms of VoIP telephony:
  - Applicatn layer VoIP (OTT VoIP) (e.g. 'Skype')
  - IP layer VoIP (e.g. VoLTE in '4G' telephony)

Lecture 9 COMP28512 8

MANCHESTER 1824
The University of Manchester

Voice over '4G'

- LTE standard supports only IP packet switching.
- 4 ways of handling voice are being discussed
- Voice over LTE (VoLTE)
  - Voice service delivered as data flows within the LTE data flow.
  - No need for legacy Circuit Switch voice networks to be maintained.
- Circuit-switched fallback (CSFB)
  - LTE just provides data services,
  - Voice calls fall back to circuit switched domain.
- Simultaneous voice and LTE (SVLTE)
  - Phone works simultaneously in LTE & circuit switched modes.
  - LTE mode for data & circuit switched mode for voice.
  - Distinction exists only in the phone.
- 'Over-the-top' (OTT) content services
  - Use VoIP apps like Skype to provide voice over LTE IP network.
  - Voice is still main revenue source, so this is not likely.

Lecture 9 COMP28512 9

MANCHESTER 1824
The University of Manchester

VoLTE

- VoLTE is in the future,
- 4G must be able to smoothly handover to 3G or 2G to maintain quality under adverse conditions.
- Demand for voice calls today has led LTE carriers to introduce 'CSFB' as a stopgap measure.
- For voice calls, LTE smartphones fall back to 2G or 3G networks.

Lecture 9 COMP28512 10

MANCHESTER 1824
The University of Manchester

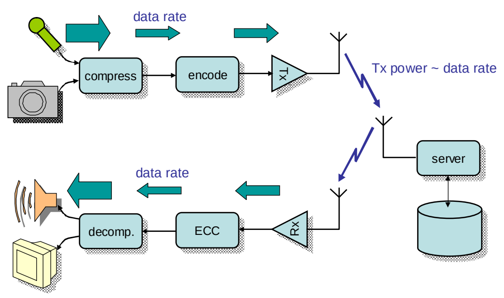
Media comms issues

- Sending to server:
  - capture the image, sound, …
  - compress to reduce data size
  - add redundancy to enable error correction
  - transmit data to server
- Receiving from server:
  - receive data from server
  - detect & correct transmission errors
  - interpolate or mask un-correctable errors
  - decompress media information
  - display or play result

Lecture 9 COMP28512 11

MANCHESTER 1824
The University of Manchester

General scheme



Lecture 9 COMP28512 12

MANCHESTER 1824 The University of Manchester

Optimizing power

- Trade-offs:
  - image/audio quality
    - better => more data
  - power required to compress data
    - more complex algorithms require more power
  - power required to transmit data
    - constant energy/bit, so more compression saves power
  - Tx power can be saved by error correction
    - better ECC allows lower energy/bit
  - additional data for error correction
    - ECC requires redundancy, so more bits must be sent
  - power required for error correction
    - again, more complex algorithms require more power

Lecture 9 COMP28512 13

MANCHESTER 1824 The University of Manchester

Coding multimedia signals

- Increase capacity of radio channels
  - data compression reduces bit-rate
  - Error Corr Coding (ECC) increases bit-rate!!
  - ECC reduces transmit power needed,
  - Decreases interference with neighbouring cells
- Minimise effect of transmission errors
  - use ECC to recover from errors
  - block interleave & similar schemes to map burst errors (normal with radio comms) into evenly distributed single errors
  - there will still be occasional un-correctable errors
    - use interpolation to minimize quality loss
    - requires careful choice of coding technique

Lecture 9 COMP28512 14

MANCHESTER 1824 The University of Manchester

Dynamic control

- Radio signals are affected by
  - reflections off buildings
  - movement, Doppler shift, antenna orientation, etc.
  - mobile & variable interference sources
  - Near-far problem
- Use feedback to maintain functionality
- Measure correctable error-rate at receiver
- Use feedback to control:
  - Tx power (e.g. transmit with power $\propto$ 1/receive-power)
  - ECC redundancy
  - Compression ratio (& hence quality)

Lecture 9 COMP28512 15

MANCHESTER 1824 The University of Manchester

Reminder about 3G-CDMA

- Each bit from coded signal is 'spread', e.g. to produce:

  1 = 1 0 1 0 1 1 0 0 0 1 0 1 0 1 0 1 0 1 1 1 0 1 0 1 0 1 0 1 0 1
  0 = 0 1 0 1 0 0 1 1 1 0 1 0 1 0 1 0 1 0 0 0 1 0 1 0 1 0 1 0 1 0

- Each 'bit' becomes a pseudo-random sequence of 'chips'
  - Transmitted at high chip-rate – needs wide bandwidth.
- Can recover orig signal if pseudo-random sequnce is known.
  - Otherwise transmission will be heard as noise.
  - All users transmit at the same time in same frequency band.
  - But they all use a different sequence
  - Receivers can recover each bit by a cross-correlation process.

Lecture 9 COMP28512 16

MANCHESTER 1824 The University of Manchester

Advantages of CDMA over 2G-GSM

1. TDMA must strictly synchronize transmission times of all users to ensure that they occupy the correct time-slot.
   - Must have 'guard-times' between slots which decrease spectral efficiency.
   - CDMA eliminates this timing synchronization, increasing spectral efficiency.
2. CDMA allows a flexible allocation of resources
   - No strict limit to the number of users (soft capacity limit)
3. With '2G', channels not being actively used are just wasted.
   - e.g. when a person is listening & not speaking.
   - With CDMA, 'silent' channels do not use any resources
   - Makes the CDMA channel usable by more users.

Lecture 9 COMP28512 17

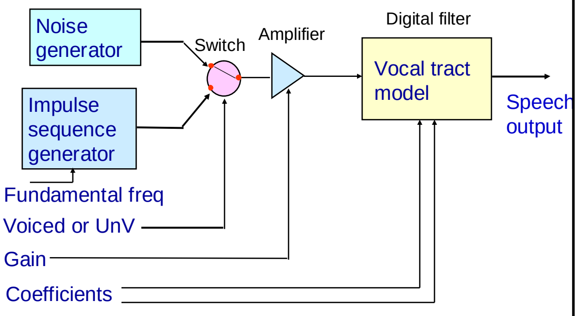
MANCHESTER 1824 The University of Manchester

Power control with CDMA

- Main disadvantage of CDMA is that power control is a critical factor & difficult to achieve,
  - Rejection of 'other' signals is only partial.
  - Unwanted nearby signals may overwhelm wanted signals from far away.
  - The 'Near-Far' problem

Lecture 9 COMP28512 18

Remember LPC-10 decoder



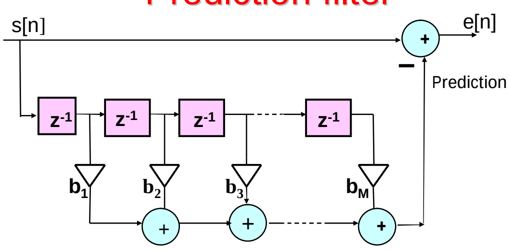
Lecture 9 COMP28512 19

LPC-10 encoder

- How are the coeffs b1, b2, ..., bM calculated at the encoder?
- Need new set of coeffs for each block of speech: $\{s[n]\}_{1,N}$
- Code & transmit e[n] & coeffs b_1, b_2, ..., b_M for each block.
- Typically M=10.

Lecture 9 COMP28512 20

Prediction filter



- $e[n] = s[n] - b_1 s[n-1] - b2s[n-2] - \ldots - b_M s[n-M]$
- For each block of speech $\{s[n]\}_{1,N}$ choose b_1, b_2, … b_M to minimise:

$$E = \sum_{n=1}^{\infty} (e[n])^2$$

Lecture 9 COMP28512 21

Finding b_1, b_2, … b_M

$$E = \sum_{n=1}^{\infty} (e[n])^2 = \sum_{n=1}^{\infty} (s[n] - b_1 s[n-1] - \ldots - b_M s[n-M])^2$$

$$\frac{\partial E}{\partial b_1} = \sum_{n=1}^{\infty} (s[n] - b_1 s[n-1] - \ldots - b_M s[n-M])\, s[n-1]$$

$$= \quad C_{10} - C_{11} b_1 - C_{12} b_2 - \ldots - C_{1M} b_M$$

$$\text{where} \quad C_{ij} = \sum_{n=1}^{\infty} s[n-i] s[n-j]$$

$$\text{Similarly,} \quad \frac{\partial E}{\partial b_2} = \quad C_{20} - C_{21} b_1 - C_{22} b_2 - \ldots - C_{2M} b_M$$

$$\frac{\partial E}{\partial b_3} = \quad C_{30} - C_{31} b_1 - C_{32} b_2 - \ldots - C_{3M} b_M$$

and so on

Lecture 9 COMP28512 22

Finding b_1, b_2, … b_M

$$E = \sum_{n=1}^{\infty} (e[n])^2 = \sum_{n=1}^{\infty} (s[n] - b_1 s[n-1] - \ldots - b_M s[n-M])^2$$

$$\frac{\partial E}{\partial b_1} = \sum_{n=1}^{\infty} (s[n] - b_1 s[n-1] - \ldots - b_M s[n-M])\, s[n-1]$$

$$= \quad C_{10} - C_{11} b_1 - C_{12} b_2 - \ldots - C_{1M} b_M$$

$$\text{where} \quad C_{ij} = \sum_{n=1}^{\infty} s[n-i] s[n-j]$$

$$\text{Similarly,} \quad \frac{\partial E}{\partial b_2} = \quad C_{20} - C_{21} b_1 - C_{22} b_2 - \ldots - C_{2M} b_M$$

$$\frac{\partial E}{\partial b_3} = \quad C_{30} - C_{31} b_1 - C_{32} b_2 - \ldots - C_{3M} b_M$$

and so on

Lecture 9 COMP28512 23

Finding b_1, b_2, … b_M (cont)

$$\text{Setting} \quad \frac{\partial E}{\partial b_1} = 0, \quad \frac{\partial E}{\partial b_2} = 0, \quad \ldots, \quad \frac{\partial E}{\partial b_M} = 0 \quad \text{we get :}$$

$$C_{10} = C_{11} b_1 + C_{12} b_2 + \ldots + C_{1M} b_M$$
$$C_{20} = C_{21} b_1 + C_{22} b_2 + \ldots + C_{2M} b_M$$
$$C_{30} = C_{31} b_1 + C_{32} b_2 + \ldots + C_{3M} b_M \quad \text{and so on}$$

$$\begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{M0} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1M} \\ C_{21} & C_{22} & \cdots & C_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ C_{M1} & C_{M2} & \cdots & C_{MM} \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}$$

$\underline{c} = [C] \times \underline{b}$

$\therefore \underline{b} = [C]^{-1} \times \underline{c}$

That's it!!

Lecture 9 COMP28512 24

MANCHESTER 1824
The University of Manchester

In MATLAB

```
for i=1:M
  c(i)=0;
  for n=1:N, if n>i, c(i) = c(i)+s(n)*s(n-i); end;  end;
  for j=1:M
    C(i,j)=0;
    for n = j+1 : j+N
      if n > i && n < N+i+1, C(i,j) = C(i,j)+s(n-i)*s(n-j); end;
    end; % n loop
  end; % j loop
end; % i loop
b = inv(C)*c';
```

Lecture 9 COMP28512 25

MANCHESTER 1824
The University of Manchester

Voiced & Unvoiced speech

•Voiced speech (vowels):

- Resonances (formants) which change as person speaks.
- These determine the vowel sound.

• Unvoiced speech (consonants):

- Vocal cords do not vibrate.
- Turbulent air flow produces "hissing" sound.
- Vocal tract excitation is random noise-like signal.

Lecture 9 COMP28512 26

MANCHESTER 1824
The University of Manchester

Effect of prediction at the transmitter

• With correctly adapted coeffs, subtracting prediction at the transmitter removes resonances (formants).

• Remaining 'prediction error' (or 'residual') signal {e[n]} becomes high-pass filtered excitation signal:

  • periodic series of pulses (voiced),
  or
  • spectrally white random signal (unvoiced).

Lecture 9 COMP28512 27

MANCHESTER 1824
The University of Manchester

Vector-quantisation of e[n]

• Instead of transmitting e[n] sample-by-sample, send several samples at once as a 'vector'.
• Store frequently occurring patterns in code-books at transmitter & receiver.
• Transmitter chooses pattern closest to the one it needs to transmit, & just sends its code-book index.
• This is 'code-book' quantisation.
• Idea like this is used in Code excited LPC (CELP).

Lecture 9 COMP28512 28

MANCHESTER 1824
The University of Manchester

'Comfort noise'

• In a 2-way telephone conversation each person may be listening or waiting about 60% of the time.
• Discontinuous transmission (DTX) is an option for not transmitting 'silence'.
  – Saves transmission power but receiver's phone may sound 'dead.'
  – No background noise heard.
• So receiver inserts some artificial background noise.
• Needs 'voice activity detector' (VAD) at transmitter.
  – Determines when talker is 'silent'
  – Characterises the background noise by some basic measurements.
  – Transmits these measurements (e.g. power) using very few bits.
• Allows receiver to synthesise 'comfort noise' that sounds approximately like the background noise at the transmitter.

Lecture 9 COMP28512 29

MANCHESTER 1824
The University of Manchester

Conclusions

• Streaming multi-media can use:
  – buffering to try to maintain a real time stream
  – interleaving to allow interpolation when packets are lost
• Dynamic control of power & FEC is vital in mobile systems
• Important for 3G-CDMA because of 'near-far' problem.
  – CDMA has nice intrinsic features for voice channels
• 4G – LTE is all IP & voice technology not yet decided.
  – OTT-VoIP is a possible but unlikely candidate.
• LPC speech encoder explained.
  – Also Vector-Quantisation as used in CELP
• Comfort noise mentioned – but does 3G-CDMA need it?

Lecture 9 COMP28512 30