University of Manchester
CS3282: Digital Communications '06
Section 1: Introduction and overview

1.1. Overview of the course

This course is concerned with the requirements and limitations of digital transmission as used for fixed and mobile telephony, wired and wireless computer networks, data storage & digital broadcasting. This is a vast and rapidly advancing subject with many aspects that are new and only recently introduced into textbooks, though often these new aspects are based on older ideas seen in a new light. To be able to appreciate the technology currently available and participate in its development there are some fundamental ideas that must be understood, a large number of technical terms and acronyms that must become familiar and some conceptualising skills that must be acquired. Perhaps the most important of these skills is the ability to visualise signals in two different ways at the same time; i.e. in the time-domain and in the frequency-domain. Experienced communications experts do this apparently without thinking. To learn some of these fundamental ideas, terms and skills is the purpose or our course.

The course begins by defining its scope with respect to applications in telephony and computer networks. The term "physical layer" is explained and the nature of wired and wireless channels over which communication takes place is discussed. Although we intend to addresses mainly the digital transmitter and digital receiver and real channel issues arising in digital communications, we will look at other aspects higher up the chain of “layers” of “protocols” governing telephone and data network communications. In many ways the digital transmitter is a digital-to-analogue converter converting the stream of binary digits (obtained possibly by digitising analogue speech) into an analogue signal suitable for the channel in question. Similarly, the receiver is a type of analogue-to-digital converter. To understand digital transmission and the demands of the analogue-digital conversion process, we must have grounding in the relationship between the shape of an analogue waveform in the time-domain and its frequency spectrum. Each binary digit will be represented by a segment of such an analogue waveform. So Section 2 of this course is a brief revision on Fourier transforms and spectral analysis.

Section 3 considers the storage and transmission of real time speech, music and video, and discusses the need for bit-rate compression. The basic ideas behind some important bit-rate compression techniques are studied, and the implementation of bit-rate compression in mobile telephony is considered. The requirements and current importance of ‘voice over IP’ (VoIP) are discussed.

Section 4 introduces the concept of digital transmission at base-band. This occurs when analogue wave-shapes generated at the transmitter are conveyed directly along a wire (or wires coupled by transformers) without the need to modulate a carrier. Very often the analogue transmission has a frequency range starting at, or close to, zero Hz, and a defined upper frequency. The simplest base-band signal representing digital data is probably a series of rectangular "pulses". Such pulses could be transmitted along an ordinary wire, though with ‘shaping’ to reduce the otherwise infinite bandwidth. A pulse is a variation in voltage say from 0 to 5 and back to 0 volts.

With ‘unipolar’ binary signalling, the presence of such a pulse could represent a logic ‘1’ with the voltage remaining zero to signify a logic ‘0’. With ‘bipolar’ binary signalling, instead of remaining at zero volts to represent ‘logic 0’, a ‘negative pulse’ would be used. This would normally be the negative of the pulse used to represent logic ‘1’, i.e. a variation from zero volts to say –5 volts and

back again. A series of unipolar or bipolar pulses, being just variations in a real directly measurable voltage, is an example of a real base-band signal. The signalling is said to be 'binary' when there are only two possible types of pulse: e.g. +5 and 0 volts for unipolar, or +5 and –5 volts for bipolar. We could have multi-level pulses, e.g. pulses of amplitude 0, 1, 2, 3, 4, 5, 6, 7 volts to represent numbers 0,1,2,3,4,5,6,7 instead of just 0 and 1, and this is no longer binary signalling.

Data connections between computers, to form a wired 'local area network' (LAN) for example, can be achieved with real base-band signalling when the distances between machines are not large, the bandwidths of the wires can be considered very large and the wires are not shared with any other transmissions such as telephone speech. Although speech is traditionally transmitted at base-band over local (home to local exchange) telephone lines, base-band transmission of pulses over these telephone lines is complicated by the fact that such telephone lines are not guaranteed to allow frequencies below about 300 Hz to pass. The "DC" component of a signal is normally lost on a local analogue telephone line, but we can shape the pulses in such a way that they do not require the zero to 300 Hz bandwidth and are therefore not affected by this loss.

Section 4 explains the meaning of 'asynchronous' transmission as commonly used for base-band signalling over short distances for text (RS232) and 'synchronous' transmission which is more efficient and suitable for higher bit-rates. As analogue pulse shapes somewhat different from the straightforward rectangular pulses mentioned above are often used for base-band transmission, some examples of these 'PCM waveform types' are presented. The effects of noise added to the received analogue waveform by the channel and receiver circuitry are analysed and ways of controlling the effect of this noise on bit-errors are discussed.

Section 5 further considers the effect of noise on bit-errors at the receiver, and introduces the concept of 'matched filtering' to minimise the effect of the noise under certain assumptions.

The frequency bandwidths of analogue pulses used to convey digital information must be adapted to the capacity of the channel. The channel will have a limited bandwidth. Rectangular pulses have infinite bandwidths and are therefore totally unsuitable. Simply rounding the sharp edges of rectangular pulses does not help much either. A more radical form of 'pulse shaping' is required which, in principle, causes each pulse to 'ring on' forever once it has started. Clearly, this 'ringing on' risks the possibility of bit-errors due to 'inter-symbol' interference. Fortunately, it is possible to eliminate 'inter-symbol' interference even when pulses do run into each other, but this can only be done when the number of pulses per second (i.e. the symbol rate) is not greater than twice the bandwidth of the pulse in Hz.

With binary signalling, each pulse represents one bit. Therefore with a pulse-shape carefully chosen to eliminate inter-symbol interference, binary signalling can achieve a 'bandwidth efficiency' of up to 2 bits/second per Hz at base-band. This is an important result established in Section 6 which concerns pulse shaping. The effects of the channel gain and phase characteristics on digital transmission are also considered in Section 6 and the need for 'equalisation' as well as matched filtering and pulse shaping is demonstrated. Traditional ways of optimising the performance of digital transmission systems using matched filtering, pulse shaping & equalisation are then considered.

The digital transmission techniques considered in this course have led, in recent years, to phenomenal advances in the speed and efficiency of digital transmission, as exemplified by the increases in bit-rates achievable over domestic telephone lines. The fact that the speeds of inexpensive data modems can now approach 56kb/s over links originally designed for telephone

voice is remarkable enough. The provision of 'broad-band' network access over the same telephone lines via ADSL is an even more remarkable achievement, and we must add to this the continuing developments of wireless telephony and the still emerging field of wireless computer networks. Broadcasting has also seen significant developments with the emergence of digital transmission for sound and video. In view of these advances, it is useful to have some idea of what is possible in theory; i.e. given the bandwidth of a channel and the signal-to-noise ratio achievable, what bit-rate is achievable given the need to maintain an acceptably low bit-error rate? The Shannon-Hartley Law, presented in Section 7, gives an answer to this question.

Many analogue and digital communication systems, both wired and wireless, involve the use of a sinusoidal signal of fixed amplitude, frequency and phase as a 'carrier'. This conveys no information itself, but it may be 'modulated' by a 'base-band' signal which does represent data. The modulation process can cause the base-band signal to modify the amplitude, frequency and/or phase of the carrier in such a way that these changes can be detected at the receiver. The advantage of having a carrier is that the range of frequencies generated can be shifted up and placed in a frequency band suitable for the channel and the equipment being used, in such as way as to avoid clashing with other transmissions.

All the ideas mentioned so far in this introduction can be understood in terms of real base-band signals. However most of these ideas may be generalised to carrier modulated signals. This is the purpose of Section 8 which starts by revising concepts of amplitude, frequency and phase shift keying (ASK, FSK and PSK). These may be thought of as 'single carrier' digital modulation schemes. Two types of digital receiver may be considered for demodulating any of these transmissions: non-coherent and coherent receivers. Non-coherent receivers are generally simpler than coherent receivers since coherent receivers require the generation of a timing waveform which is synchronised exactly in frequency and phase with the received carrier. Although this synchronisation can require quite complicated circuitry and synchronisation techniques it is well worthwhile.

The use of a coherent detector allows us to use a 'vector-modulator' to produce more efficient forms of single carrier amplitude, frequency and phase shift keying (ASK, FSK and PSK). To understand these, we introduce the concept of a 'complex base-band' signal which simply comprises two independent real base-band channels, one of which modulates the carrier $\cos(2\pi f_C t)$ while the other modulates $\sin(2\pi f_C t)$. The $\cos(2\pi f_C t)$ and $\sin(2\pi f_C t)$ carriers are at the same frequency f_C, and when the two separately modulated carriers are added together, the result is a more efficient 'single carrier' modulated signal than is obtained by modulating just $\cos(2\pi f_C t)$ or just $\sin(2\pi f_C t)$. In fact twice the bit-rate may be carried since the properties of $\cos(2\pi f_C t)$ and $\sin(2\pi f_C t)$, i.e. their 'orthogonality' when coherent detection is used, allow each of the two base-band modulating signals to be detected separately as if the other was not present. So we can transmit two independent channels at the same carrier frequency. We refer to $\cos(2\pi f_C t)$ as the 'in phase' carrier component and $\sin(2\pi f_C t)$ as the 'in quadrature' carrier component. The use of vector-modulation with PSK, i.e. applying binary PSK independently to $\cos(2\pi f_C t)$ and $\sin(2\pi f_C t)$ and summing the result, produces a widely used technique known as 'quadrature PSK' or QPSK.

With binary signalling, each symbol or pulse represents one bit. With a suitable pulse-shape, binary signalling can achieve a bandwidth efficiency of up to 2 bits/second per Hz at base-band. Since multiplication by a sinusoidal carrier doubles the bandwidth of a real base band signal by introducing upper and lower sidebands, up to a maximum of 1 bit/second per Hz with can be achieved with single carrier modulation. However, with vector-modulation, two channels each

giving 1 bit/second per Hz are possible, thus bringing the overall achievable bandwidth efficiency back to 2 bits/second per Hz. This is at the expense of the additional complexity of coherent detection. Whereas binary PSK can achieve a maximum of 1 bit/second per Hz, QPSK can achieve 2 bit/second per Hz because of the 'in phase' and 'quadrature' components of the transmission.

With a band-width efficiency of 2 bit/second per Hz, a computer modem could achieve a maximum of about 6 kb/s over a 300-3400 kHz domestic telephone link. This is the maximum achievable with binary signalling and is little more than one tenth of what we know to be possible. A similar calculation for the bandwidth-efficiency of binary FSK is a little more complicated, but if we take the frequency spacing used by MSK, we may expect to achieve a maximum of slightly less than 2 bit/second per Hz.

To increase the band-width efficiency over what can be achieved with binary signalling we must use multi-level modulation schemes where each symbol represents more than one bit. Section 9 deals with multi-level carrier modulation & demodulation and the combined use of multi-level ASK and PSK to produce efficient schemes such as QAM and APK which are widely used in data modems to achieve bit-rates approaching 56kb/s.

The explosion of interest in wireless (radio) communications, including mobile phones and wireless computer networks has produced applications for bandwidth-efficient high speed digital transmission techniques. As well as being efficient in their use of bandwidth, these techniques must also be well suited to radio channels subject to frequency-selective fading. Frequency-selective fading occurs due to reflections of radio signals from buildings and walls interfering with each other and cancelling each other out at certain frequencies. It becomes more and more serious as we widen the band-widths of channels to achieve higher bit-rates. The final part of Section 8 introduces the very latest 'multi-carrier' modulation schemes (OFDM) and their advantages for wireless computer networks. OFDM is also used for digital radio and TV broadcasting. As the name implies, multi-carrier techniques use more than one carrier frequency. Digital radio and TV broadcasting uses 1024 carriers and wireless LANs use 64. Fortunately the modulation process is achieved on all 1024 or 64 carriers by means of a single fast Fourier transform FFT computation. The use of the FFT and a "cyclic" extension has a profound effect on the pulse-shaping, matched filtering and equalisation problems arising with single carrier systems. In fact the pulse-shaping and matched filtering become unnecessary and the equalisation process is greatly simplified.

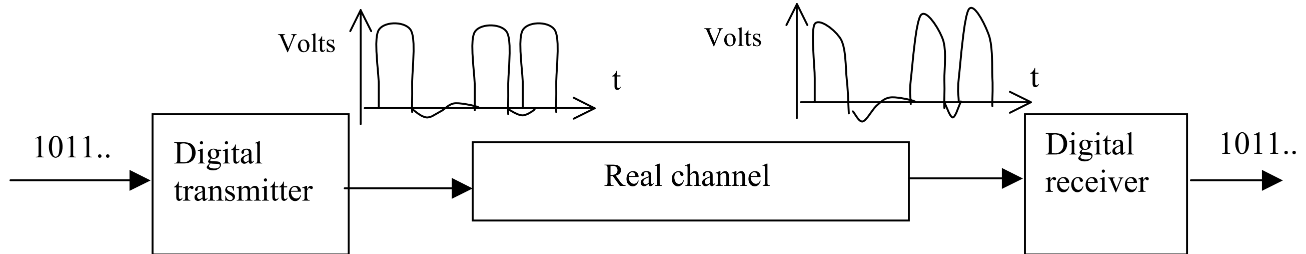
Finally, we discuss multiple access techniques for digital communication by radio. Sharing a given radio bandwidth efficiently between users, e.g. mobile phone users, is possible in many different ways, including code division multiple access which is used in the USA currently for 2G mobile telephony and will be used for 3rd generation mobile phones world-wide.

Having surveyed the whole course, the remainder of this section introduces some background information about digital transmission channels as used by computer and telephone networks.

1.2. Digital transmission channels

It is generally assumed in Computer Science that we can send sequences of binary digits such as 1011011110000111001… , over digital transmission channels and receive them exactly, or perhaps with a few bit-errors, some distance away. This is achieved by placing a *digital transmitter* and a *digital receiver* at the ends of a *real channel* which is invariably analogue in nature.

In some ways the digital transmitter is a 'digital-to-analogue converter'. It converts the bit-stream to a sequence of pulses suitable for the real channel. Similarly, the digital receiver may be considered to be an 'analogue to digital converter' which takes the pulses received from the real channel, usually distorted somewhat, and coverts them back to a bit-stream. This process is illustrated in the block-diagram below:



In traditional telephony, the *real channel* may be wire or cable, optical fibre, a radio channel, an infrared or ultrasound link or perhaps a combination of several of these types. Such channels are also used for transmitting broadcast material for TV and radio and the data generated by computer networks and the Internet.

Another form of transmission channel is obtained by magnetic and optical recording devices and their corresponding read-back devices (e.g. magnetic disks, tapes, CDROMs, DVDs, Zip-drives, etc.). The similarities between the technical problems of storage and transmission are as striking as the differences between them. For example, efficient bandwidth utilisation and error concealment raise similar issues in both cases whereas overall delay and cost requirements are considerably different.

Where the need for apparently "instantaneous" transmission ("real time" communication as required, for example, to hold a normal telephone conversation) is not present, cost savings can be enormous using magnetic/optical storage and retrieval rather than on-line transmissions. For example, carrying a 30 gigabyte hard disk by train from Manchester to Liverpool would constitute a channel of about 80 megabits/second ($30 \times 8 \times 10^9 \div 50 \div 60$ as the journey takes about 50 minutes) at a cost of about 1 pence per gigabit. Consider the alternative of sending the data over a telephone line using a 56kb/s modem?

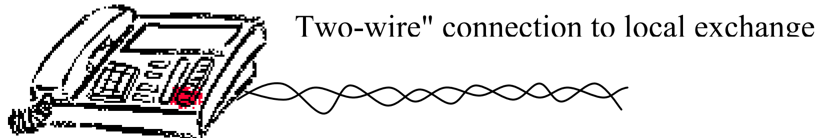
It will become apparent that the performance of a digital transmission link is governed by three main factors: (a) usable channel bandwidth, (b) received noise & (c) channel in-band frequency characteristics. Factors (a) and (b) are generally constant whereas factor (c) is highly variable especially for mobile communication systems. Mobility introduces a 4th factor which is a 'Doppler' frequency shift which occurs due to the movement of the transmitter with respect to the receiver or vice-versa.

We will discuss the design of digital transmitters and receivers with reference to the bandwidth and frequency-response of the channel and the received noise characteristics. Fundamental limitations in what can be achieved in a given situation will be established. This is referred to as the "physical layer" of a communication system which, in the field of computing and computer networks, is the lowest layer of the "OSI reference model".

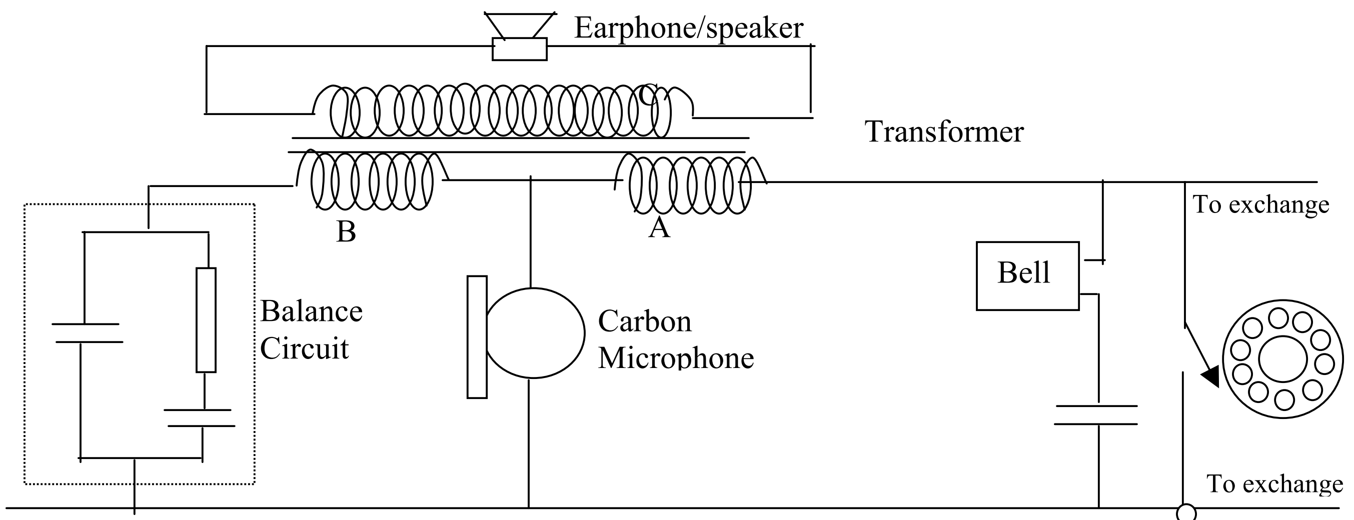
1.3.Telephone networks:

It is useful to have a brief discussion about *plain old-fashioned land-based telephony* (POTS). Land-based telephone links from the local exchange to the home or office, the "local loops", are still predominantly analogue, though this may change quickly with the widespread availability and reducing cost of ISDN and ADSL.

<u>1.3.1. The terminal</u>: The familiar telephone handset at home is more complicated than it looks. It is linked to the local exchange by just 2 wires which must provide signalling (dialling and ringing) and then carry a 2-way conversation. To ring the bell, the exchange sends a 17 Hz AC (75 volt) signal.



Dialling was for many years by "loop disconnect " signalling (largely out-dated now) and is now mainly by multi-tone (DTMF) signalling. In principle, the circuitry in a domestic telephone is as shown below, though these days more modern components are used:



The microphone was originally a carbon microphone with compressed carbon granules acting as variable resistor. A battery (in the local exchange) provides quiescent current of about 20 mA and speech causes resistance and therefore the current to vary. The earphone was originally a "rocking armature" type of speaker, using an electromagnet. To achieve a 2-way conversation, a transformer in the handset reduces side-tone (i.e. the talker hearing his own voice much louder than that of the distant talker) by detecting the talker's signal and generating an almost equal and opposite signal to cancel most of it out. Referring to the diagram above, any current changes in A are mirrored by equal and opposite changes in B and sum to zero in C. The balance circuit is intended to approximately match the impedance of the line to the local exchange. Signals from the local exchange pass through A and B in the same direction and hence sum at C to provide the earphone signal.

<u>1.3.2. Bandwidth</u>: Although the bandwidth of speech extends much wider, it was found many years ago that intelligibility and a large degree of naturalness is preserved if the bandwidth is limited to frequencies below about 4 kHz. For traditional and some technical reasons, the bandwidth of "telephone quality" speech was and still is defined to be about 300Hz to 3.4kHz. Originally, this allowed speech to be conveyed, in analogue form, in 4kHz wide frequency-division multiplexed

channels, on high capacity "trunk" wires or point-to-point radio links between exchanges. The 4kHz channel bandwidth gave some margin for imperfect filtering. Later, the same 300 to 3.4 kHz bandwidth was adopted for digital transmission, allowing the speech to be sampled at 8 kHz. The lower limit of 300 Hz, originally needed mainly because of the use of transformers and electro-mechanical switches, has been retained perhaps unnecessarily.

Virtually all POTs 'exchange to exchange' transmission is now digital and radio-based mobile telephony is now predominantly digital throughout. Protocols for telephone communication exist as they do for computer networks and it is meaningful to talk about the physical layer as the basic mechanism by which binary digits from digitised speech are transmitted over a real channel from one location (e.g. a local exchange) to another.

1.3.3: The Channels

As mentioned above, connections between domestic users and the local exchange are still normally via line-pairs carrying analogue signals. A single pair of wires, i.e. one complete circuit, is provided for each user which means that this circuit must carry 2-way information. This does not mean that one of the two wires carries a "go" signal with 'earth' to complete the loop with the other wire, with earth, carrying a "return" signal. It is theoretically possible to do this but not a good idea. Originally the links were often bare copper wires on wooden telegraph poles. More commonly, they are "twisted pairs" in underground cables carrying many similar pairs. Twisting reduces "cross-talk" arising from currents induced magnetically from nearby pairs. A local exchange will serve thousands of telephone customers within perhaps a one or two mile radius within a town or city, though over greater distances in the countryside. The exchange must be able to connect each customer to the local exchange of a called person who may be tens, hundreds or thousands of miles away. Clearly there cannot be a separate wire for each call, and multiplexing many channels along a much smaller number of high capacity links is required. Examples of the types of 'exchange to exchange' channels used are as follows:

1. **Twisted Pairs**: - As well as providing links from the local exchanges to the subscribers, twisted pairs were also commonly used in the past for single channel exchange-to-exchange '2-way' trunk lines. To reduce attenuation at speech frequencies (0-4kHz), loading coils (88mH) were placed at 1.8km spacing. Later, low capacitance twisted pairs were used for long distance transmission of up to 60 one way channels (4kHz spacing) by analogue frequency division multiplexing. This of course required the '2-wire' (single circuit) communication with the domestic user to be converted to a '4-wire' link with separate 'go' and 'return' circuits "Repeaters" amplifying the signals were placed at suitable spacing. A high precision cable design was needed to achieve a usable bandwidth of 250 kHz.

2. **Coaxial Cables:** - These were (and still are in some places) used as long distance trunk lines carrying very large numbers of multiplexed channels. Their bandwidths are typically 0.3 to 60 MHz. Before the introduction of digital transmission, 10,800 channels would be carried using frequency division multiplexing, with repeaters at 1.8km.

3. **Microwave Radio Links**: - Line of sight transmission at carrier frequencies of 1-12GHz. Transmitting and receiving dishes typically placed on 100 metre towers spaced 40 to 50km apart. Typically about 300,000 speech channels (or equivalent) are transmitted between two stations. TV channels are transmitted by this means also (one TV channel ≅ 3000 telephone speech channels). Satellites also provide microwave radio link capacity.

4. **Optical Fibre**: - Now the predominant technology for exchange-to-exchange links. Thin glass wire carrying pulses of light, have the advantages: -
    a. Low loss (< 1 dB/km)
    b. Large bandwidth
    c. Small in size, low weight, and inexpensive
    d. No electromagnetic interference.
    e. Small cross-talk, high security
    f. Resistance to ageing and temperature variations

    Repeater spacing: 2-7km for multimode fibre and 15-50km for monomode.

1.3.4: Analogue Transmission: - Frequency-division multiplexed channels of bandwidth up to 30 MHz can carry up to 6000 (4kHz bandwidth) telephone links. "Two-wire" domestic links carrying bi-directional analogue speech are split at the local exchange into separate go/return channels, i.e. "4-wire links" each band-limited to about 4 kHz. The channels are then multiplexed together to be carried efficiently between exchanges by high capacity channels (using cable, wireless links, etc.) The following frequency-division multiplexing hierarchy was commonly used before digital transmission became predominant

a. *Group*: - 12 go (or 12 return) channels transmitted simultaneously in the band 60kHz-108kHz. Each channel is allocated a 4kHz band. Single side-band modulation was used to place the audio signal in its transmission band.
b. *Supergroup*: - 5 groups separated by fdm into 312-552kHz.
c. *Hypergroup*: - 15 or 16 supergroups i.e. 900 or 960 channels.

1.3.5: Digital transmission: - Although digital links to the home are becoming more commonplace (ADSL, ISDN etc.), in most cases digitisation of telephone speech takes place at the local exchange, the "last mile" or so to the domestic user remaining "2-wire" and analogue. Again '2-wire' to '4-wire' conversion must take place at the local exchange so that the "send" channel can be digitised. Speech channels are sampled at 8kHz with 8 bits/sample to give 64kbits/sec. Pulse code modulation (PCM) then produces a 64kb/s bit-stream representing the analogue signal, and this is multiplexed with other digitised channels and transmitted uni-directionally along digital transmission channels. The capacity of a higher bit-rate digital exchange-to-exchange link is shared by time-division multiplexing (TDM) whereby bits from one channel are interleaved between those of others. Typical digital transmission systems exist with the following and now higher bit-rates.

100Mbits/s ------------ 1400 channels
400Mbits/s ------------ 5760 channels
1600Mbits/s ----------- 23040 channels

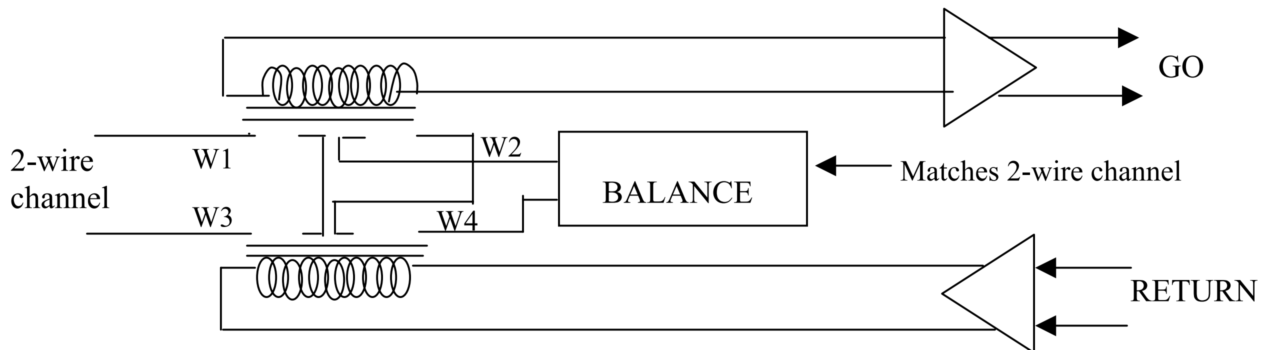
. . The following time-division multiplexing (tdm) hierarchy is standard in the UK:

a. *Group*: - 32 channels (30 available for speech, 2 used for synchronisation & signalling). Bit-rate 2.048Mbits/s.
b. *Supergroup*: - 4 groups.
c. *Hypergroup*: - 16 supergroups. (1920 channels at 139.264Mbits/s)

The introduction of digital transmission techniques brought great advances and economies in exchange to exchange signalling. For example, twisted pair links originally designed to carry a single bi-directional analogue speech channel (with loading coils at about 2km spacing) could be redeployed for one 'time division multiplexed' group (with the loading coils replaced by repeaters performing digital regeneration). The introduction of digital transmission therefore allowed these same lines to be used to transmit, *one way*, 32 PCM channels by time-division multiplexing (tdm). *Coaxial cables* were, until the advent of optical fibre, commonly used for high capacity digital transmission using tdm with repeater spacing of about 1.5km for 400Mbits/sec links (5760 channels

i.e. 3 hyper-groups). Regenerative repeaters were used to preserve the pulse shape for digital transmissions. Also equalisation was employed to improve the gain and phase response of the link.

1.3.6: Two/four wire conversion: - Multiplexed analogue and digital transmission channels use separate "go" and "return" links. They are one-way only. As we have discussed, subscribers' lines use a single 2-wire path for both "send" and "receive". Hence "two-to-four" line and "four-to-two" line conversion is needed at the local exchange. This is provided by a circuit called a "hybrid".



1.3.7: Wireless telephony**:** There are two basic types of wireless telephone: cordless and "cellular" mobile. Both types have developed rapidly in recent years, but the greatest developments have been with cellular mobile phones. These have gone through two "generations" and are entering a third. The first generation was analogue, as were non-cellular mobile phone systems pre-dating the cellular concept. The current (second) generation is digital and focused largely on voice transmission. The third generation will integrate digitised voice, data, the Internet, e-mail and lots of other services traditionally provided by computer networks. Comparing with POTs, it may be considered that the major differences are that

(i) wireless links replace the 'exchange-to-user' link (or 'local loop') , i.e. the local loop.
(ii) the radio medium is shared and
(iii) the 'base-stations' with which mobile phones communicate and which, in many ways, fulfil the role of small "local exchanges" (very local) can "hand off" customers to other base stations when customers walk or drive down the street, or even when a base-station becomes overloaded with customers.

Otherwise base-station to base-station communication takes place over the same sort of links and channels as are used for POTs, though these are advancing rapidly in capacity and availability as well. The "local loop" wireless links are not single channels per user as with POTs, and separate radio channels are used for the 'go' and 'return' speech paths, and also for the dialling and signalling. We refer to the "down-link" from base-station to mobile and the "up-link" from mobile to base-station. The radio medium is shared by various forms of multiplexing (e.g. fdm, tdm and cdma) and also by strategically placing base-stations such that the same frequency bands may be re-used by base-stations that are far enough apart as not to interfere with each other. This is the "cellular" concept and is sometimes referred to as "space-division multiplexing". A simplistic view of the cellular concept has nice hexagonal areas of land (cells) with a base-station in the centre of each cell serving it in a frequency band which is different from that used by neighbouring cells. In practice, base-stations are often placed on tall buildings such as churches and rely on shielding by buildings to minimise interference. The big problem with wireless mobile to base-station links is the effect of multi-path propagation since the radio transmission from a given base-station or a given mobile user will be reflected by buildings and thus reach its intended target by a multitude of

different routes. In fact, direct "line-of-sight" paths between base-stations and mobiles are rare in large cities with tall buildings. The different routes introduce different phase shifts, and so when the reflected signals reach their target some of them will add in phase and reinforce each other, and some will be out of phase and can cancel each other out. The cancellation is referred to as "fading", and it will be frequency dependent, i.e. the same phase differences will cause reinforcement at some frequencies and cancellation at other frequencies. Further, if the mobile is moving, the nature of the fading will change with time, often very rapidly. If this is not complicated enough, there will also be a frequency shift due to the Doppler effect. The effect of multi-path propagation is to make the gain and phase response of the radio channel non-flat and highly variable. The channel acts like a filter whose gain and phase response vary. If we restrict transmissions to a relatively narrow band of frequencies, the gain and phase differences across this narrow band may not be too serious, and the main problem will be "flat fading", i.e. the gain and phase-delay across the whole band being affected more or less equally due to cancellation. However, if we wish to increase the bit-rate of our transmissions by using a wider radio-frequency band, "frequency selective" fading will become a problem where some frequencies within the wider band will be severely attenuated while others are enhanced.

The "coherence bandwidth", B_C Hz say, of the radio channel is the largest bandwidth we can use without having to worry about frequency selective fading. For a channel whose bandwidth is less than B_C Hz, any fading may be considered "flat fading". B_C is often assumed to be proportional to $1/\sigma_t$ where σ_t is the "r.m.s.delay-spread" in seconds, i.e. a measure of the spread of delays due to multipath propagation. The value of σ_t , and hence B_C , depends on the environment , for example the spacing of buildings within a city. A typical value for a city of B_C = 30 kHz would allow an analogue mobile telephone system with 30 kHz channels to work without an equaliser, whereas a 900 MHz GSM system with 200 kHz bandwidth channels would definitely require equalisation. An "equaliser" is a filter or frequency-selective amplifier whose frequency-response is, in principle, the inverse of that of the channel over the selected frequency band, so that the effects of frequency selective fading are reversed and the channel then appears to have a constant gain and phase-delay across all frequencies.

Fading affects the phase-response as well as the gain-response of a radio channel. A non-linear phase response means that the phase delay is not the same for all frequencies so that some frequency components of a signal will be delayed by different amounts of time than others. Such a phase-response can seriously alter the shape of a signalling pulse and thus cause "inter-symbol interference" (ISI) when it runs into the next pulse to arrive. Reductions in the gain-response due to fading cause a smaller signal to be received, and therefore noise picked up by the receiving antenna will have a greater effect and produce more bit-errors. These are fundamental problems with the physical layer as used by wireless telephony, and they also arise wireless LANs. Similar problems occur with wired links and ordinary telephone lines have non-ideal gain and phase responses which must be compensated for, usually by an equaliser, if used by high speed data modems. However, these problems are not so dramatic or rapidly changing as they are with radio. Ways of dealing with these problems for wired and wireless digital transmission will be addressed in later sections.

1.3.8. Advantages of digital voice networks in telephony

1. Ease of multiplexing: - TDM (time-division multiplexing) equipment is simpler and less expensive than FDM (frequency-division multiplexing) equipment even when the cost of digitising speech is taken into account.

2. Ease of signalling: - Control information (e.g. on/off hook, dialling, coin deposits, charging) is inherently digital and can therefore be transmitted in exactly the same way as digitised speech.
3. Use of modern technology: - May be applied to switching, multiplexing and signal processing. Digital I/Cs are easier to manufacture than analogue components especially as the latter are usually required to be linear. Advantages of digital signal processing (replacing analogue filters, echo-cancellers etc.) are:
    a. *Reproducibility*: - i.e. digital circuits can be copied exactly and are not subject to temperature effects or ageing.
    b. *Programmability*: - i.e. systems are flexible and can be changed.
    c. *Time-sharing*: - i.e. the same circuit can process multiple signals.
    d. *Automatic testing*: - i.e. testing can be routinely performed.
    e. *Versatility*: - i.e. the same circuit can process multiple signals.
4. Integration of transmission and switching: - The multiplexing operations of a transmission system may be easily integrated into the switching equipment.
5. Operability at low signal-to-noise-ratios: - Analogue noise and interference is most noticeable during speech pauses or when the amplitude is low. Noise in digital systems is mostly quantisation noise produced at the A-to-D converter and since companding and adaptive gain control (mirrored at the receiver) may be used to encode low level signals with the same signal to quantisation noise ratio as high level signals (approx.), we do not need an excessively high ratio for high level signals just to ensure that the ratio for low level signals is acceptable. Also, crosstalk in analogue systems is particularly annoying as it tends to be intelligible. Even if crosstalk in digital systems is bad enough to cause bit-errors, the resulting noise will be unintelligible and therefore not as disturbing.
6. Signal regeneration possible: - Distortion in analogue transmission systems cannot be corrected. In a digital system, the probability of transmission errors can be made arbitrarily small by inserting regenerative repeaters, which reconstruct the original pulse shape at intermediate points in the transmission link.
7. Accommodation of other services: - A transmission link can be totally indifferent to the nature of the traffic it carries - it is just a bit-stream. Hence data, video, music etc., can be accommodated in a totally integrated system.
8. Performance monitoring: - Possible, for example by recording parity errors.
9. Ease of encryption: - Scrambling and unscrambling of the transmitted bit stream for security is much easier.

1.3.9: Disadvantages of digital voice networks in telephony

1. PCM transmission: - Requires a number of pulses (typically 8) to be transmitted for each sample. This could mean that the bandwidth required would be multiplied by at least this number, as compared with analogue, since the sampling rate, f_s must be greater than twice the signal bandwidth and a transmission bandwidth of $1/(2T)=f_s/2$Hz is needed for normal pulse transmission.
2. Analogue/digital conversion needed: - Although the cost of this conversion is usually offset by savings in other parts of the system.
3. Need for time synchronisation: - Whenever digital information is transmitted, a timing reference or 'clock' is needed for deciding when to sample the waveform, to decide what bit is being transmitted. It is inconvenient, usually, to send this timing reference separately: it is usually included in the digital transmission and this inclusion enforces constraints as to how the data is transmitted (e.g. long strings of zeros or ones must be disallowed if the receiver clock is likely to lose synchronisation). More subtle problems arise with time diversion multiplexing since clocks used to transmit signals from different parts of the telephone network may not be

accurately synchronised. Time division multiplexing signals whose clock rates are slightly different causes some problems requiring "guard-bands" between channels for tolerance.

4. Topologically restricted multiplexing: synchronisation of clocks for TDM becomes more and more difficult as the geographical locations of the transmitting stations is allowed to become wider and wider (because of the increasing range of delays). For this reason, TDM has been used primarily in applications where all the information sources are centrally located and single multiplexer controls the occurrence and assignment of time slots - although more sophisticated synchronisation techniques are being introduced.
5. Incompatibility with existing analogue facilities: - Until the telephone system becomes totally digital, many awkward and expensive interfaces will be needed.

1.4. Computer networks:

1.4.1 Wired networks:

The equivalent of telephony's "local loop" for wired computer networks is probably the "local area network" (LAN) which, these days, usually uses some form of Ethernet to interconnect computers (or "hosts") within a large building or office. The name Ethernet (or "IEEE 802.3") specifies the type of wire or cabling used for the LAN, the packet-switched digital transmission method (which is a form of time-division multiplexing though not as rigidly timed as in telephony) and the MAC protocols used to share the capacity between all users. Originally Ethernet transmissions were carried by coaxial cables (like TV aerial cabling) snaking round a building forming a single "bus" onto which every host was able to connect and compete for its share of the transmission capacity. These days, "twisted pair" wire connections extending radially from a "hub" or "switch" are more common. This is known as "10Base-T" Ethernet standing for 10Mb/s, base-band signalling, twisted pairs. As with the POTs local loop, a single twisted pair connects each host bi-directionally to the hub or switch, which must be no more than 100 meters away. Regenerative repeaters may be used for longer links.

Within a large building, exactly the same types of twisted pair connections can be used for the telephones and the computer LAN. Because of the short distances involved, almost the simplest possible form of base-band digital transmission can take place along these lines using more or less normal pulses with "Manchester coding". A 'hub' is simply a way of interconnecting hosts as though they were on a common bus. It is a small box with a large number of connectors one for each computer. With earlier forms of Ethernet (e.g. 10BASE-2 or "thin Ethernet") where all hosts are connected to a single coaxial cable acting as a "bus", all users had to compete for access to this bus according to a set of rules known as "medium access control" MAC protocols (see later). Although '10Base-T' Ethernet does not have the same bus structure the use of a hub would mean that essentially the same "collision sensing multiple access"(CSMA) protocols as were used with 'thin Ethernet' are still used. These allow "collision detection" (CSMA/CD) to determine when two devices are trying to transmit at the same time and "collision avoidance" (CSMA/CA) by sensing the channel first and waiting until it is clear. Just as though they were hosts connected to a common bus, all computers connected to a hub are said to be in the same "collision domain", i.e. when one host transmits a frame (i.e. a packet of bits as generated by a host at its data link layer) any other host connected to the hub will cause a collision if it tries to transmit at the same time.

The existence of a device (i.e. the hub) with a separate radial connection to each host makes CSMA to some extent unnecessary. Therefore hubs which simply connect all their radial links together are becoming obsolete. They are being replaced by 'switches' which perform essentially the same task as hubs but incorporate buffering, delays and fast electronics to multiplex all the received packets in

a more efficient way. This is "switched Ethernet". Whereas repeaters and hubs operate entirely at the physical layer, doing no more than amplifying and interconnecting voltage waveforms representing frames of bits (without bothering about what information these frames contain), switches are able to examine the frames, to be precise the header information, to determine which computer to route the frame to. Switches are therefore said to operate at the data link layer (see next section) rather than the physical layer. Details about this and more recent and faster forms of Ethernet may be found in the textbook by Tanenbaum, for example.

1.4.2: Bridges, routers and gateways:

We now consider how a wired LAN may be connected to the outside world. The connection could be to another wired LAN, or to the Internet. This is of course possible via a modem over a telephone line, but this restricts us to a single point-to-point link and relying an a single ISP. A "bridge" is a type of switch which is used to interconnect two or more LANs. The two devices are almost identical and both use the same header information inserted in each frame by data link layer to make the right connections. However, whereas the switch is used to interconnect a single LAN where all hosts use the same type of Ethernet, the bridge is able to communicate with different types of LAN which may not even be Ethernet based. Routers are very different from switches and bridges, and need to look more deeply at the content of each frame; i.e. a router needs to examine the "packet header" inserted at the TCP/IP "network layer" to form an "IP packet". The packet header indicates the destination and according to this information the router chooses the most appropriate of its output lines, e.g. a link to London, to send the packet on its way towards its destination. The router is, in some ways, the computer network equivalent of a local telephone exchange or a cellular base-station and in many cases uses the same type of channels and digital transmission techniques. In fact the router may share exactly the same channels with telephone transmissions.

Gateways are able to connect devices that use completely different protocols. For example they can connect TCP/IP computer networks to telephone networks so that voice over IP (VoIP) users can make ordinary telephone calls. Gateways are also used to allow TCP/IP frames to be conveyed by 'asynchronous transfer mode" (ATM) networks since there is much spare capacity on such optical-fibre based networks (designed for packetised telephone and real time traffic) though the packet sizes (TCP/IP frames and ATM cells) and headers are completely different and translation and repacketising is necessary.

1.4.3. Protocols, Layers and the OSI and TCP/IP Reference Models

A *protocol* is an agreement between communicating devices on how communication is to proceed. An analogy is the behaviour of people when they meet, especially at formal occasions. For example, "How do you do, I am Mr. Smith", offer hand, "How do you do Mr. Smith, I am Miss Jones", shake hand, remove hand after a suitable number of shakes, etc. Getting things wrong can be catastrophic at formal occasions and also in electronic communications. As different nationalities have different customs; bowing, kissing hands, etc. there has to be understanding and tolerance on the protocols to be adopted. The same is true for electronic communications, and international committees have been set up to define the required protocols.

Electronic communication protocols are defined in *layers*. A well-known description of computer-to-computer communication is known as the *OSI Reference Model*. *OSI* stands for "Open Systems Interconnection", a term invented by an international committee called the "International Standards Organisation" in about 1983. An "open system" is simply a computer or other device open for

communicating with other devices. The concept of a "reference model" is that we can relate each part of a computer-to-computer communication link to a clearly defined and easily understood diagram. A "model" is supposed to ease understanding.

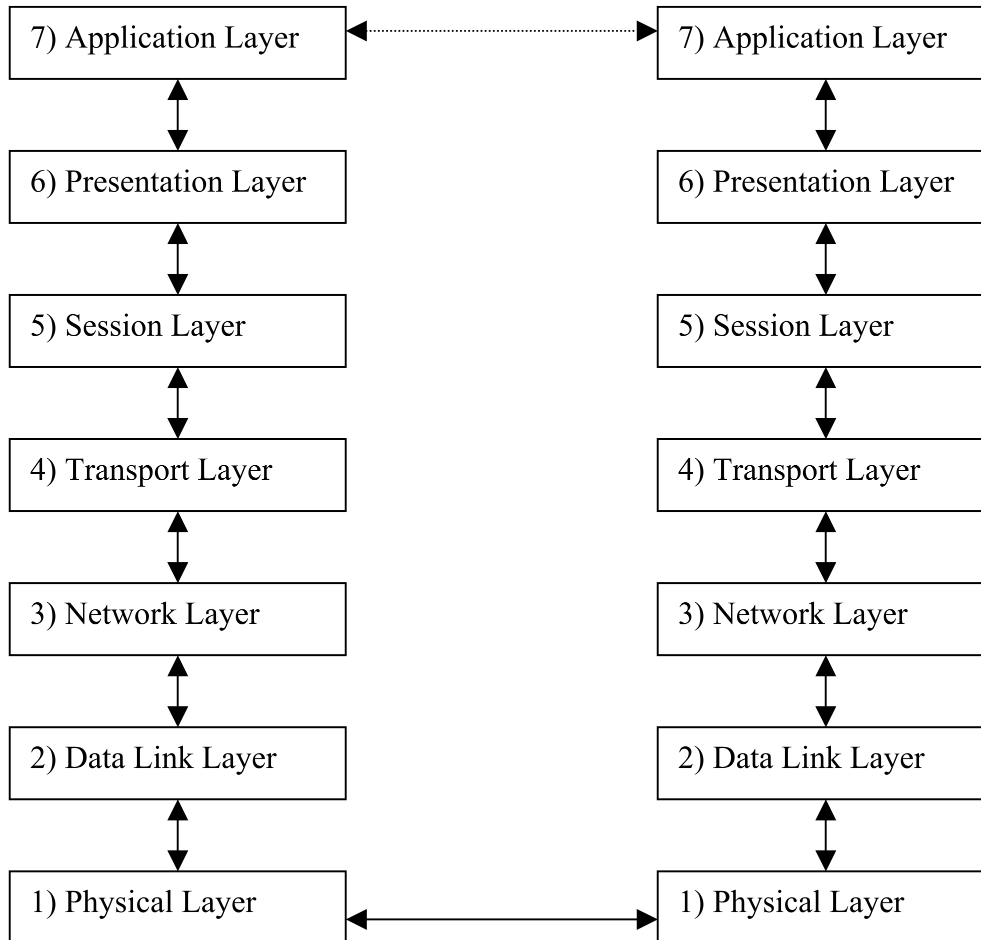


Figure 1.1: The "7-layer OSI reference model"

The *OSI reference model* is actually rather (and some say unnecessarily) complicated with seven layers. The lowest layer is the ***physical layer*** which controls the method by which raw binary digits are transmitted over the real channel and received. With wireless LANs the physical layer is divided into two sub-layers: the "physical layer convergence procedure" (PLCP) sub-layer and, below this, the "physical medium dependent" (PMD) sub-layer. The physical layer must do its best to ensure that when the transmitting computer sends a binary "1" the receiver gets a single bit of value 1, and not a zero bit or perhaps two bits instead of one. It is inevitable that errors will occur in this process from time to time and the bit-error rate (BER), i.e. the number of bits in error per second should be estimated. It is also useful to know whether these errors will tend to occur in bursts (i.e. a lot of errors close together) or whether they will be fairly evenly spread out in time. The synchronisation of data rates (clock timings) is an important aspect of the physical layer also. The physical layer can accommodate various types of transmission including "packetised" data where binary information is collected together into blocks or "frames" consisting of typically a few hundred or a few thousand bytes. At the physical layer, packets are often referred to as "frames". With Ethernet and wireless LANs, the physical layer must be able to generate a "clear channel

assessment" (CCA) flag to be used by the MAC sub-layer of the data-link layer (see next) to determine whether the channel is clear for a transmission.

The ***data-link layer*** aims to use the raw data transmission facility (with errors) offered by the physical layer to achieve a computer-to-computer link that is essentially error free. Here the data to be transmitted is supplied by the "network layer" as a packet. This packet must be split into "data frames" suitable to be sent via the physical layer. Since the physical layer is designed simply to accept and transmit streams of bits regardless of their meaning or structure, the data link layer must create frame boundaries and include special bit patterns before and after the data to allow the corresponding physical layer in the receiving system to recognise the frame boundaries. The effect of physical layer channel errors may be reduced by including extra "channel coding" bits to allow "forward error correction" or at least the "detection" of errors. Introducing "parity" bits allowing error detection and "Hamming" codes allowing error correction are simple examples of error coding methods that may be employed by the data-link layer. Another task of the data-link layer is to generate and send out "acknowledgement" frames when required to do so, and to recognise acknowledgement frames when they are received. Acknowledgement frames are sent to tell the other computer that its data-frame has been received. When an acknowledgement frame is received, we know a data-frame we have just sent has been received by the other computer. Finally and most importantly for "broadcast" networks such as bus-based wired local area networks (LANs) using "Ethernet" and wireless LANs which share the resources of a single channel, the data link-layer has a "sub-layer" known as the ***"medium access control" (MAC) sub-layer***. The MAC sub-layer offers a whole set of protocols for allowing different devices to have access, at different times, to the shared channel without devices constantly clashing or hogging the resources.

The ***network layer*** adds information to the data it receives from the transport layer above it to determine how the routers in a network will transmit the information. This information is in the form of a "packet-header" which contains messages for the routers and can also enable charging (i.e. billing) by network operators. A possibility here is to specify different routes for different frames. Perhaps there could be fast routes for some frames and slower routes for others.

The ***transport layer*** splits up a block of data from the session layer above into units small enough to be handled by the network layer, and passes these down to the network layer with additional information (another header) to ensure that if the units take different transmission paths and arrive out of order at the receiver, (a common occurrence with packet transmission systems) they can be put back into the correct order. The transport layer may divide the data from the session layer into smaller units each of which is specified to use a different network connection (for high speed). Or it may do the opposite, i.e. multiplex several transport connections on to the same network connection (for convenience and reduced cost). The transport layer is most often required to provide an error free point-to-point channel for use by the session layer above it.

The ***session layer*** provides the protocols for users to log-on to different machines, enter passwords etc., transfer files to and from these different machines and generally control the dialogues that take place.

The ***presentation layer*** is for standardising the coding conventions for data transmitted between computers. Layers below the presentation layer just think of the data as ones and zeros. The presentation layer will convert, for example, different coding schemes for alpha-numeric characters (such as ASCII, EBCDIC and Unicode) and different signed multiple-byte number representations to a common standardised form to make it possible for computers which use these different

representations to communicate without confusion. Examples of different number representations in common uses are:

- "two's complement" and "offset binary" representations for signed integers,
- "low byte first" and "high byte first" representations for 16-bit integers and
- many different forms of floating point numbers.

The top layer is the ***application layer*** which has a variety of functions including that of emulating a given terminal so that the operation of the computer screen and keyboard are matched to what is required by the communicating system. Other tasks concern the standardisation of file naming conventions and the transmission of electronic mail in an acceptable form.

The **"TCP/IP Reference Model"** is similar to the 'seven layer OSI' model and pre-dates it. It was first defined in 1974 for connecting together multiple networks in a seamless way and in it lays the origin of the Internet. It is less rigidly defined than the OSI model, and may be thought of as having four layers.

The ***internet layer*** defines a standard packet format and transmission protocol called "IP" (Internet Protocol). The *internet layer* sends the packets to where they are supposed to go. This function is similar to that of the *OSI network layer*.

Above the "TCP/IP internet layer" is the ***TCP/IP transport layer*** which is similar to the *OSI transport layer*, except that there are two different end-to-end protocols possible:-

i) **TCP** (Transmission control protocol): which is designed to be a reliable error free protocol with acknowledgement and retransmission facilities. The data to be sent is fragmented into suitably sized messages (data blocks) which are passed down one by one to the internet layer for transmission. If a suitable acknowledgement is not received for each message within a given time the appropriate message block is retransmitted.
ii) **UDP** (User Datagram Protocol): A "fire and forget" protocol which sends messages as for TCP but does not have acknowledgement/retransmission facilities.

The TCP/IP reference model does not have session or presentation layers. The ***Application Layer*** contains all the higher level protocols, i.e. those in OSI levels 5, 6 and 7. Among these are TELNET, FTP, SMTP, DNS, NNTP, HTTP and many others.

The ***Host-to-Network Layer*** is not closely defined and varies from host to host and from network-to-network. Clearly it must include the function of the OSI physical layer.

A comparison of the two models is given below: -

Layer	OSI Reference Model	TCP/IP Reference Model
7 6 5	Application Layer Presentation Layer Session Layer	Application Layer
4	Transport Layer	Transport Layer
3	Network Layer	Internet Layer
2 1	Data Link Layer Physical Layer	Host-to-Network Layer

Table 1.1: - Comparison between OSI Reference Model and TCP/IP Reference Model

1.4.4: Wireless computer networks

It is possible that a revolution is about the happen, or is happening, in the field of computer networks that will rival that we have witnessed in telephony. It may indeed be the same revolution which will result in the two technologies finally merging into one. There are certainly two distinct starting points and currently two identifiably different approaches to the delivery of speech multimedia, data and Internet access to the customer; i.e.

(a) third generation telephony, including UMTS, which will be strongly connection oriented and based on a traditional multiple access approach (CDMA) in highly expensive licensed radio frequency bands (remember the recent auctions of bandwidth),

(b) wireless LANs which are packet-switched, capable of operating in 'ad-hoc' mode without infrastructure such as the base-stations required by cellular mobile telephony. There is no need for users to sign contracts with or make payment to operating companies for access to their base-stations. And finally, the wireless LAN devices operate in the 2.4 to 2.48 GHz (Industrial Scientific & Medical) or 5.17 to 5.27 GHz radio frequency bands which are free, freely available and currently unused except for microwave ovens, garage door openers and a small number of similar devices. The user can have Internet access via connections to high capacity wired networks, as with mobile telephony, but without having to pay for the wireless access to a base-station or computer acting as an "access point".

Wireless LANs can certainly replace or augment wired LANs for more convenient data communication and access to the Internet. However, they may also be able to combine this operation with real time communication thus providing convenient access to telephone services as well as data networks.

Just as the IEEE 802.3 standard known as "Ethernet" is almost universally used for wired LANs, there is an IEEE standard known as "IEEE 802.11" sometimes referred to as "Wi-fi" which is likely to be predominant in the field of wireless LANs. There are many variations of the basic standard, the most important ones currently being IEEE802.11a, IEEE 802.11b and IEEE802.11g. A competing European standard known as Hiperlan, with two variations numbered 1 and 2, also exists though many people consider its future to be uncertain. Hiperlan 1 is already obsolete, though Hiperlan2 which is very similar to IEEE 802.11a may survive. An attempt to standardise the interconnection by radio of computer peripherals, PDAs, printers etc. over very short ranges of up

to about 10 metres led to the development of "Bluetooth". The fact that Bluetooth uses the same 2.40-2.48 GHz frequency band as IEEE802.11b and that it is non-compatible with IEEE802.11 is a serious problem. Bluetooth is capable of wiping out IEEE802.11b communication over its short range. The IEEE 802.11 committees are considering this problem and most likely will propose a new standard to rival Bluetooth.

The physical layer for wireless LANs, i.e. the radio channel, presents similar problems to those discussed earlier for mobile telephony, e.g. multi-path fading and additive noise. However, the radio frequencies are higher, distances are generally less, the required data-rates are much higher and the multiple access techniques used are completely different. With wireless LANs, multiple access to the radio channel by different devices is most commonly achieved by techniques which are more like those used by Ethernet than fdm, tdm or cdma as used in telephony. In fact there are two possible modes of communication between IEEE 802.11 devices: PCF mode and DCF mode. Both modes can co-exist in the same wireless LAN.

PCF stands for "point co-ordination function" and PCF mode is possible where there is a central IEEE802.11 device, say a computer, which can act as a central controller (or base-station) by informing all other devices when they are allowed to transmit or receive data. It does this by periodically sending control packets or "beacons" to enable or disable PCF mode and "polling" devices by sending further control packets to request data from each device. A wireless LAN with a central controller capable of fulfilling this co-ordination role is termed an "infrastructure" network and, in most cases, the central controller also provides access to the outside world, e.g. via a telephone connection, and it is then termed an "access point".

DCF stands for "distributed co-ordination function" and DCF mode does not require a central controller for determining when IEEE802.11 devices can transmit or receive. DCF mode is often referred to as "ad-hoc" mode since it allows two devices within range to communicate with each other without any planning or co-ordination by other devices. DCF mode is similar to Ethernet, though there are some important differences. Although DCF mode works without an access point it can certainly have one for access to the outside world and it can give the access point priority in its access to the radio medium. Instead of having a central controller governing access to the radio channel, DCF mode uses "carrier sense multiple access" (CSMA) as does Ethernet. A WLAN device can sample the medium and determine whether any device is currently transmitting. Collision avoidance (CSMA/CA) strategies are then employed to ensure, as far as possible, that a device transmits only when the radio channel is guaranteed to be free of other traffic. One important difference between IEEE802.11 medium access control (MAC) strategies and the Ethernet strategy is that unlike an Ethernet device, a wireless LAN device cannot listen while it is transmitting so that "collision detection" (CSMA/CD) is not possible. A further difference, and a very serious problem for wireless LANs, known as the "hidden node" problem, is that two devices A and B say, may be in range with a third device, C, but out of range with each other. If A transmits to C and B cannot sense or detect this transmission, B may start transmitting also thus causing a collision with A's message and rendering it useless. In some cases, maybe for short messages, we may choose to take a chance and allow such "hidden node" collisions to occur from time to time relying on retransmissions (with randomised time delays or back-off as normally used with Ethernet) to achieve corrected transmissions. In other cases, however, and especially with longer messages it is safer to use a "request-to send/clear-to send (RTS/CTS) protocol between devices before any of them starts a transmission. The RTS/CTS protocol requires the sending and receiving of control packets which, of course, can also collide with other hidden node transmissions They are made to very short to minimise this occurrence.

1.4.5: Original IEEE 802.11 physical layer for wireless LANs:

Having described the two main medium access control strategies available under all varieties of IEEE802.11, we now consider the digital transmission techniques used. The first IEEE802.11 standard came out in 1997 offering a bit-rate of either 1Mb/s or 2Mb/s. It operates in a "licence exempt" radio frequency band between 2.4 and 2.48 GHz which is reserved for industrial, scientific and medical (ISM) equipment including microwave ovens and garage door openers. Restrictions are imposed on the permissible power levels radiated in this ISM band and there is a statutory requirement to use "spread spectrum" techniques to spread the transmitted power over a wider spectrum than is actually needed. This reduces the average power spectral density (Watts/Hz) and makes transmissions appear noise-like to traditional receivers. The effect of the spreading is reversible by a spread-spectrum receiver which knows how the spreading was achieved.
Originally IEEE802.11 had a choice of two frequency spreading techniques: "frequency hopping" (FHSS or FH) and "direct sequence" (DSSS or DS) and two transmission bit-rates: FHSS causes a carrier to "hop" around from one frequency to another in a pseudo-random pattern which is known by the receiver. IEEE802.11 divides the available 2.4-2.48 GHz ISM band into about eighty 1MHz wide bands (this is the maximum bandwidth allowed) and hops the carrier frequency into the centres of these bands, for a "dwell time" of about 0.4 seconds each, the bands being chosen in a pseudo-random fashion. DSSS increases the bit-rate by "exclusive-or"ing each bit with a "chipping sequence" i.e. a much faster sequence of pseudo-randomly chosen bits. The faster bits are referred to as chips. With IEEE802.11, the chip sequence is the same for every bit and is well known as the "11-chip Barker sequence" which is {1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0}. So, for example, a 1MHz bit stream containing:

{…, 0, 1, 1, …}

would become an 11 MHz chip-stream containing:
{…, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,   0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1,   0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, … }

The receiver generates a succession of Barker chipping sequences:
{…, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,   1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,   1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, … }

and 'exclusive-or's with the received 11 MHz chip-stream to obtain:
{ …, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,   1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,   1, 1, 1, 1, 1, 1, 1, 1, 1, 1, … }

Down-sampling this to 1 Mb/s reconstructs the original data. If the received bit-stream were not correctly synchronised, e.g. delayed by one chip for some reason, we would obtain:
{ …, ?, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, ….. }

where the number of ones and zeros per bit is close to being equal. The receiver can therefore try to achieve synchronisation by delaying the chip-stream to maximise the number of equal chips per bit. The number of chips per bit (eleven here) is the "spreading ratio".

IEEE802.11 frames have three parts: preamble, header, and payload. The preamble is a sequence of bits, typically ..101010101…, followed by a "start frame delimiter". The preamble allows the receiver to recognise and synchronise with a transmission, and to know exactly when the header will start. The header indicates the number of bytes in the payload and the data-rate of the payload. The variable length payload then follows with some error detection bits. With original IEEE802.11, the preamble & header were always sent at 1Mb/s and there was a choice between 1Mb/s and 2Mb/s for the payload.

The FHSS form of IEEE802.11 used 80 bits for its pre-amble, 32 bits for its header and two-level Gaussian frequency shift keying (GFSK) to modulate its hopping carrier at 1Mb/s. To achieve 2 Mb/s for the payload, four-level GFSK was used.

The original DSSS form of IEEE802.11 used 144 bits for its pre-amble, 48 bits for its header, always at 1Mb/s. The bit-rate was increased to 11 Mb/s by DSSS with a Barker chipping sequence and used to modulate a carrier in the 2.4-2.45 GHz band using binary differential phase-shift keying (binary DPSK). To achieve 2 Mb/s for the payload, i.e. 22 Mb/s after spreading, quaternary DSPK (or differential DPSK) was used instead of binary DPSK. With a bandwidth efficiency of one DPSK or DQPSK symbol per 2 Hz, the required bandwidth is 22 MHz in both cases. Since DPSK & QPSK could operate at closer to 1 symbol per Hz, this seems a generous allocation of frequency and it means that if there is other wireless LAN operating in the same 2.4-2.48 GHz band, (about 80 MHz wide) its carrier frequency must be at least 22 MHz away. Note that the bit-rate for the preamble & header is always fixed at 1 MHz, and that only the payload bit-rate can vary.

1.4.6 Physical layer for the latest versions of IEEE802.11:

The bit-rates achievable by the original versions of IEEE802.11 were considered too low and, in 1999, two new standards emerged: IEEE802.11a and IEEE802.11b. These new standards use essentially the same MAC sub-layer protocols as the original, but offer higher bit-rates by using more efficient and elaborate modulation schemes in some cases just for the payload.

IEEE802.11a, was proposed for use in the 5.17 to 5.8 GHz frequency band which offers about 600 MHz of unlicensed bandwidth and is referred to as "the 5 GHz band". It is based on a "multi-carrier" modulation technique known as "orthogonal frequency division multiplexing" (OFDM). The term "multiplexing" here is possibly misleading as OFDM is definitely a modulation technique and not primarily a way of combining several different wireless LAN transmissions. It is used because a single transmission is shared out among a large number of different carriers. OFDM is used in broadcasting TV and digital radio (DAB) where the number of carriers is 1024. With IEEEE_802.11a there are 64 carriers, and a bit-rate of 54Mb/s is achievable. The principles of OFDM as used for wireless LANs will be explained in a subsequent section.

IEEE802.11b operates in the same 2.4 to 2.48 GHz "Industrial scientific & medical (ISM)" band used by the original standard, adopts the same 1Mb/s data-rate and DSSS / DSPK modulation technique as the DSSS version for the preamble and header. However it is capable of using a different modulation technique for the payload to achieve up to 11 Mb/s for the payload only. IEEE802.11b products have been on the market for some time now, whereas IEEE802.11a devices are only starting to appear. The header is as before except that instead of specifying either 1 or 2 Mb/s for the payload, it can specify 1, 2, 5.5 or 11 Mb/s. To achieve 5.5 Mb/s or 11 Mb/s the same differential QPSK technique (with phase shifts of 0, $\pi/2$, π and $3\pi/2$) is used to transmit 11 chip sequences per bit at 22 M chips/second , but these sequences are no longer identical '11-chip Barker sequences'. They are varied to carry data according to a rather elaborate re-coding scheme. This technique is known as complementary code keying (CCK).

In November 2001, a new standard was defined for the 2.4-2.48 GHz ISM band which offers the advantages of IEEE 802.11a while being able to co-existing with IEEE802.11 and IEEE802.11b transmissions. This is known as IEEE 802.11g which can deliver a payload at up to 54Mb/s in a 2.4GHz band. As with IEEE802.11b, the preamble and header are transmitted at a lower bit-rate to remain compatible with earlier versions of IEEE 802.11. Since regulations specify that only

"spread spectrum" modulation techniques can be used in 2.4 GHz ISM bands, OFDM had to be re-classified as a spread spectrum technique.

1.4.7. Bluetooth:

A short range network connecting devices such as computer peripherals and PDAs is said to be a "piconet". Bluetooth is a standard for such a piconet which specifies radio access and frequency hopped single carrier modulation techniques which are similar in principle to those used by the original 1 Mb/s FHSS IEEE802.11 standard over wider distances. Bluetooth is not an IEEE standard and is not easily described in terms of the seven-layer OSI model. It operates in the 2.40-2.48 GHz band over ranges of about 10 metres .The band is divided into 79 channels each 1MHz wide and Bluetooth frequency hops its carrier with 1600 hops/second and a dwell time of 625 microseconds. There is a master unit controlling the hopping sequence used by each slave. All devices hop together. Frequency shift keying (FSK) modulation is used with a bandwidth efficiency of 1 bit/s per Hz to achieve 1Mb/s. Much of this capacity is consumed by a rather elaborate overhead of control information. The fact that Bluetooth uses the same 2.40-2.48 GHz frequency band as IEEE802.11b and that it is non-compatible with IEEE802.11 is a serious problem since Bluetooth is capable of wiping out IEEE802.11b communication over its short range. The IEEE 802.11 committees are considering this problem and most likely will propose a new standard to rival Bluetooth.

1.5 Comparing telephone networks and computer networks

There is much commonality both physically and conceptually between computer networks and telephone networks. Channels designed for speech are routinely used for digital data and to form computer network links. Also, since speech is almost always digitised for exchange-to-exchange transmission, the techniques used for digital transmission of speech and data are very similar and in many cases identical.

An important difference is that telephony is traditionally a "connection-oriented" service where one telephone is set up to be connected to another and remains connected until the call is terminated. Computer networks, however, are essentially "connection-less" and may be thought of as being modelled on the postal service rather than the old-fashioned telephone service. In this analogy, packets of information are like letters, each with an address. Each packet may take a different route, and sequences of packets may even arrive out of order. However, the packets are often sent over the same kinds of channels as are used for telephone calls.

Nowadays, defining telephone networks as "connection-oriented" and packet switched networks as "connectionless" is too simplistic (note the words "traditionally" and "essentially"). There has been much effort to merge these two approaches, for example by:

(a) transmitting digitised speech over packet switched networks such as the Internet (e.g. voice over IP, "VoIP"),
(b) reserving transmission capacity over packet-switched networks to make a given link between two users appear to be a connection (e.g. asynchronous transfer mode (ATM) networks).

Note also that when a user dials up his Internet Service Provider (ISP) over a domestic telephone line he is using a connection-oriented link for his packetised computer data. A special "point-to-point" protocol (PPP) is available for such data connections.

The issue with examples (a) and (b) above is "quality of service (QoS)" as required for real time speech, and, in the future, audio-visual links, between two or more telephone users.

1.6. Conclusions:

This section surveys the course and discusses some of the main applications of digital transmission referring to fixed and wireless computer networks and wired and wireless telephony.

1.7. Problems and discussion points:

1. As this is a course on digital transmission, why need analog FT?
2. Why is lowest 300Hz bandwidth lost on POTs telephones?
3. Why do we need a hybrid & what happens when it works badly?
4. What causes 'fading' in a mobile telephone channel?
5. What is meant by the 'rms delay spread'?
6. From list, what do you consider the main advantage of digital?
7. If IEEE802.11 is 'wi-fi' what is IEEE802.3 ?
8. What is the 'mac-sub-layer'?
9. Why is the 11-chip Barker sequence used with IEEE802.11b?
10. Design a telephone modem using a PC sound card.
11. What does FHSS have to do with the first nude actress?
12. What are the 'ISM' bands & what are they used for?
13. Why are there no bit-errors in emails?
14. Why is TCP/IP not ideal for VoIP?
15. What is the 'hidden node problem' & how is it solved?
16. What is the difference between CSMA/CD & CSMA/CA?
17. Why is 2-to-4 line conversion needed in the public switched telephone network and, in principle, how can this can be achieved using a "hybrid". What would be the consequences of an incorrectly matched hybrid?
18. What is the difference between echo suppression and echo cancellation?
19. Give the full titles of the following TLA's (3-letter acronyms) and FFLA's: PSTN, POTS, ISDN, UMTS, CCITT, ITU, ETSI, GSM, PCM, AWGN BER, pdf, NRZ, ACF, FDMA, TDMA, FSK, PSK, SNR, SQNR, CDMA, LAN, ATM, PAM, ISI, ISI.
20. What is meant by "time division multiplexing"(tdm)? Seven 64kb/s digitised speech channels and one 64kb/s signalling channel are to be tdm multiplexed to form a 512kb/s trunk transmission. (In practice this would be 30 speech plus 2 signalling channels giving 2048kb/s). Using digital multiplexers, demultiplexers and counters (available as integrated circuits) design sequential/combinational logic circuits for the tdm at a local exchange and demultiplexing at a receiving exchange. Assume a synchronised 512kHz clock is available at transmitter and receiver, and that the 8 channels are also synchronised to a common 64kHz clock derived from this 512kHz clock. What sort of information do you think is needed in the signalling channel? If a 512kHz synchronised clock were not available at the receiver, how could it be derived from the transmission itself?
21. From the list given, what do you feel are the three main advantages and the most important disadvantage of digital voice networks in telephony?
22. If you did not know anything about how data modems for domestic telephone lines work, which may be true, and you had to devise, in a great hurry, a simple technique for transmitting a string of bits over the available 300 to 3400 Hz bandwidth, using a PC equipped with a sound-card, suggest three different approaches which might work.

(Updated 30/1/06 BMGC)