ELEC3028 (EL334) Digital Transmission

Half of the unit:

- **Information Theory**
- **MODEM** (modulator and demodulator)

Professor Sheng Chen: Building 53, Room 4005
E-mail: sqc@ecs.soton.ac.uk

Lecture notes from: Course Office (ECS Student Services)
or Download from:
http://www.ecs.soton.ac.uk/∼sqc/EL334N/
http://www.ecs.soton.ac.uk/∼sqc/EL334/

Other half: **CODEC** (source coding, channel coding) by Prof. Lajos Hanzo

What's New (Should be Old) in Communications

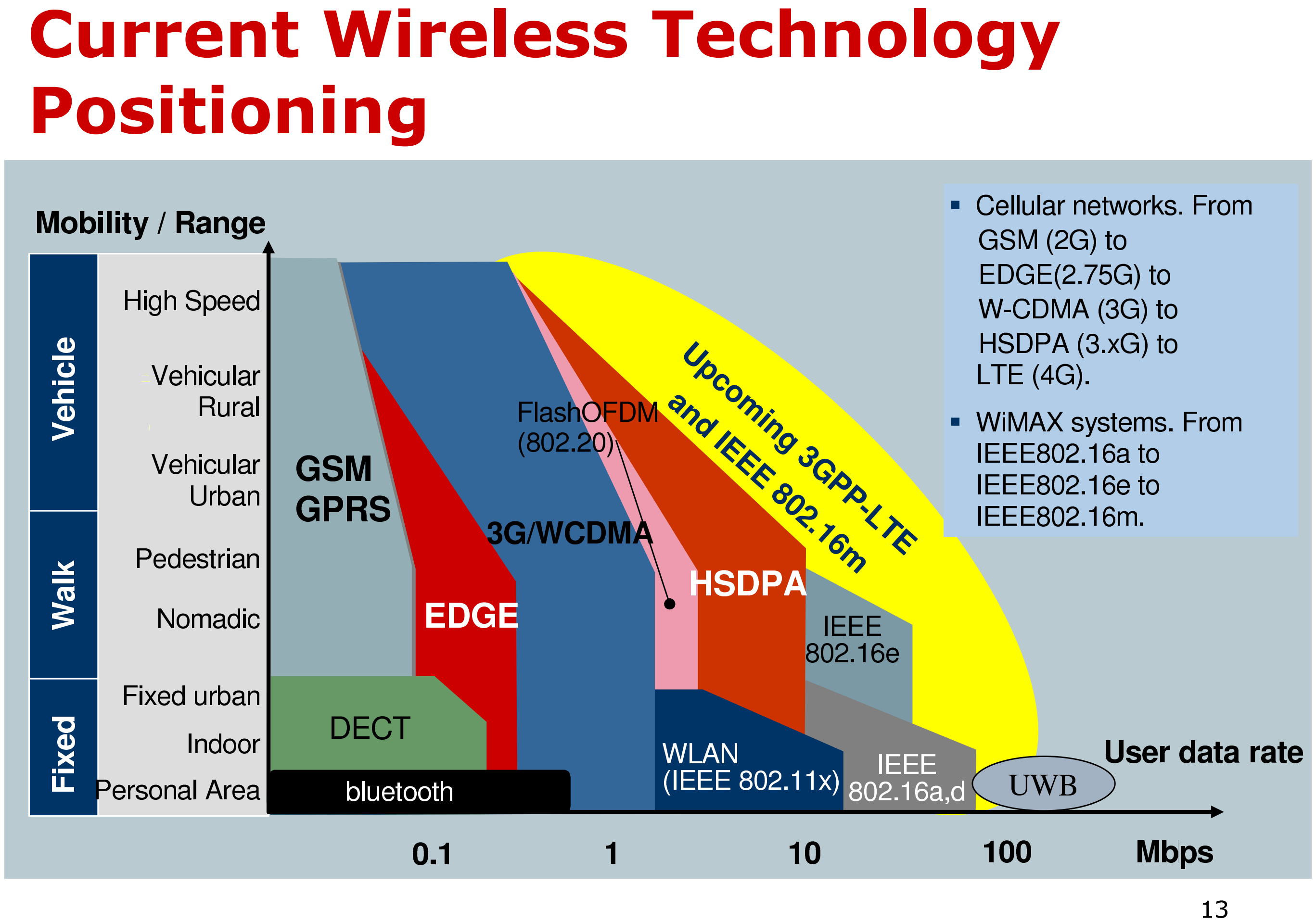
- Imagine a few scenarios:
  - In holiday, use your fancy mobile phone to take picture and send it to a friend
  - In airport waiting for boarding, switch on your laptop and go to your favourite web side
  - Or watch World Cup with your mobile phone
- Do you know these words:

  CDMA, multicarrier, OFDM, space-time processing, MIMO, turbo coding, LTE

  Broadband, WiFI, WiMAX, intelligent network, smart antenna, IP telephone
- In this introductory course, we should go through some A B C ... of **Digital Communication**

Wireless and Mobile Networks

- Past/current/future: 2G, 3G, B3G or 4G



Some improved 2G, HSCSD: high-speed circuit switched data, GPRS: general packet radio service, EDGE: enhanced data rates for GSM evolution. Also, HIPERLAN: high performance radio local area network

General Comments on Communications

- Aim of telecommunications: to communicate information between geographically separated locations via a communications channel of adequate quality (at certain rate reliably)

- The transmission will be based on digital data, which is obtained from (generally) analogue quantities by

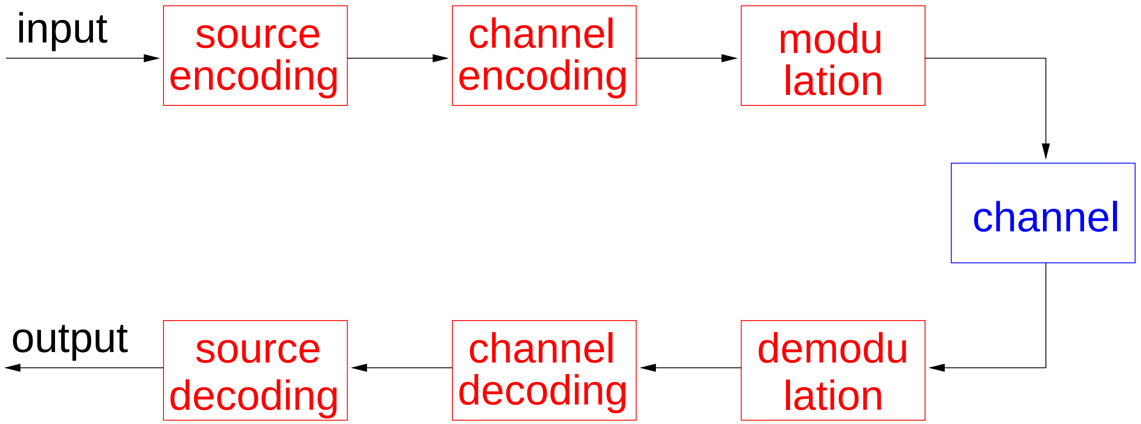
  1. sampling (Nyquist: sampling with at least twice the maximum frequency), and
  2. quantisation (introduction of quantisation noise through rounding off)

- Transmitting at certain rate requires certain spectral **bandwidth**

- Here channel means whole system, which has certain **capacity**, the maximum rate that can be used to transmit information through the system **reliably**

General Transmission Scheme

- A **digital transmission** scheme generally involves:



- Input/output are considered digital (analogue sampled/quantised)
- **CODEC**, **MODEM**, **channel** (transmission medium)
- Your 3G mobile phone, for example, contains a pair of **transmitter** and **receiver** (together called **transceiver**), consisting of a CODEC and MODEM

What is Information

- Generic question: what is **information**? How to measure it (**unit**)?
- Generic digital source is characterised by:
  - Source alphabet (message or symbol set): $m_1, m_2, \cdots, m_q$
  - Probability of occurrence (symbol probabilities): $p_1, p_2, \cdots, p_q$
    e.g. binary equiprobable source $m_1 = \mathbf{0}$ and $m_2 = \mathbf{1}$ with $p_1 = 0.5$ and $p_2 = 0.5$
  - Symbol rate (symbols/s or Hz)
  - Probabilistic interdependence of symbols (correlation of symbols, e.g. does m_i tell us nothing about m_j or something?)
- At a specific symbol interval, symbol m_i is transmitted correctly to receiver
  - What is *amount of information* conveyed from transmitter to receiver?
  - The answer:

$$I(m_i) = \log_2 \frac{1}{p_i} = -\log_2 p_i \quad \text{(bits)}$$

Concept of Information

- Forecast: tomorrow, rain in three different places:
  1. Raining season in a tropical forest
  2. Somewhere in England
  3. A desert where rarely rains

- **Information content** of an event is connected with **uncertainty** or inverse of probability. The more unexpected (smaller probability) the event is, the more information it contains

- **Information theory** (largely due to Shannon)
  – Measure of information
  – Information capacity of channel
  – coding as a means of utilising channel capacity

Shannon Limit

- We know different communication system designs achieve different performance levels and we also know system performance is always limited by the available *signal power*, the inevitable *noise* and the need to limit *bandwidth*

- What is the ultimate performance **limit** of communication systems, underlying only by the fundamental physical nature?

- Shannon's information theory addresses this question

  Shannon's theorem: If the rate of information from a source does not exceed the capacity of a communication channel, then there exists a coding technique such that the information can be transmitted over the channel with arbitrarily small probability of error, despite the presence of noise

- In 1992, two French Electronics professors developed practical *turbo coding*, which approaches Shannon limit (transmit information at capacity rate)

Information Content

- Source with independent symbols: $m_1, m_2, \cdots, m_q$, and probability of occurrence: $p_1, p_2, \cdots, p_q$

- Definition of information: amount of information in ith symbol m_i is defined by

$$I(m_i) = \log_2 \frac{1}{p_i} = -\log_2 p_i \quad \text{(bits)}$$

Note the unit of information: bits !

- Properties of information
  - Since probability $0 \le p_i \le 1$, $I(m_i) \ge 0$: information is nonnegative
  - If $p_i > p_j$, $I(m_i) < I(m_j)$: the lower the probability of a source symbol, the higher the information conveyed by it
  - $I(m_i) \to 0$ as $p_i \to 1$: symbol with probability one carries no information
  - $I(m_i) \to \infty$ as $p_i \to 0$: symbol with probability zero carries infinite amount of information (but it never occurs)

Physical Interpretation

- Information content of a symbol or message is equal to minimum number of binary digits required to encode it and, hence, has a unit of bits
  - Binary equiprobable symbols: $m_1, m_2 \leftrightarrow \mathbf{0}, \mathbf{1}$, minimum of one binary digit (one bit) is required to represent each symbol
    Equal to information content of each symbol: $I(m_1) = I(m_2) = \log_2 2 = 1$ bit
  - Four equiprobable symbols: $m_1, m_2, m_3, m_4 \leftrightarrow \mathbf{00}, \mathbf{01}, \mathbf{10}, \mathbf{11}$ minimum of two bits is required to represent each symbol
    Equal to information content of each symbol: $I(m_1) = I(m_2) = I(m_3) = I(m_4) = \log_2 4 = 2$ bits
  - In general, q equiprobable symbols m_i, $1 \leq i \leq q$, minimum number of bits to represent each symbol is $\log_2 q$
    Equal to information content of each symbol: $I(m_i) = \log_2 q$ bits

- Use $\log_2 q$ bits for each symbol is called Binary Coded Decimal
  - Equiprobable case: m_i, $1 \leq i \leq q$, are equiprobable $\Rightarrow$ BCD is good
  - Non-equiprobable case ?

Information of Memoryless Source

Source emitting a symbol sequence of length N. A memoryless Source implies that each message emitted is independent of the previous messages

Assume that N is large, so symbol m_i appears $p_i \cdot N$ times in the sequence

- Information contribution from ith symbol m_i

$$I_i = (p_i \cdot N) \cdot \log_2 \frac{1}{p_i}$$

- Total information of symbol sequence of length N

$$I_{total} = \sum_{i=1}^{q} I_i = \sum_{i=1}^{q} p_i \cdot N \cdot \log_2 \frac{1}{p_i}$$

- Average information per symbol (entropy) is

$$\frac{I_{total}}{N} = \sum_{i=1}^{q} p_i \cdot \log_2 \frac{1}{p_i} \quad \text{(bits/symbol)}$$

Entropy

- Memoryless source **entropy** is defined as the average information per symbol:

$$H = \sum_{i=1}^{q} p_i \cdot \log_2 \frac{1}{p_i} = - \sum_{i=1}^{q} p_i \cdot \log_2 p_i \qquad \text{(bits/symbol)}$$

- Source emitting at rate of R_s symbols/sec has **information rate**:

$$R = R_\mathrm{s} \cdot H \qquad \text{(bits/sec)}$$

- If each symbol is encoded by $\log_2 q$ bits, i.e. binary coded decimal, average output bit rate is $R_\mathrm{s} \cdot \log_2 q$. Note information rate R is always smaller or equal to the average output bit rate of the source[1] !

source: alphabet of q symbols

symbol rate R_s

$\log_2 q$ bits (BCD)

actual bit rate: $R_\mathrm{s} \cdot \log_2 q \geq R$

[1] Hint: $H \leq \log_2 q$

Summary

- Overview of a digital communication system: system building blocks
- Appreciation of information theory
- Information content of a symbol, properties of information
- Memoryless source with independent symbols: entropy and source information rate