CSC 407 : DATA COMMUNICATION AND NETWORK

BASICS OF DATA COMMUNICATION AND COMPUTER NETWORKING

1.0. INTRODUCTION

Computers are used to generate information. Generated information is not useful in itself. The information must be delivered to the right destination at the right time. Often information must be transmitted from one location to another. This process is called data communication. Here, we will be concerned with the hardware, software and procedures used in data communication. Office automation is based on communication and the transfer of information. Advances in communication technology, combined with rapidly evolving computer technology, have made possible much of the progress in the field. Electronic communication consists of telecommunication and data communications. Telecommunication refers to the use of telephone, telegraph, and radio or television facility to transmit information, either directly or via computer. Data communication means the transfer of data or information between computer devices.

1.1. DATA COMMUNICATION

Data communication is the active process of transporting data from one point to another. Networks are communication system designed to convey information from a point of origin to a point of destination. Note that they are communication system, not computer system. The operative word is communication, the transfer of information from one device to another. Networks come in two flavours – **local** as in local area network, which cover a small area and have a finite, relatively small number of users and **global** which cover long distance and have an unlimited number of users. Telephone networks are long network. It refers to the transmission of the digital signals over a communication channel between the transmitter and receiver computers. Communication is possible only with wired and wireless connectivity of the computers with each other. The effectiveness of a data communication system depends on three fundamental characteristics:

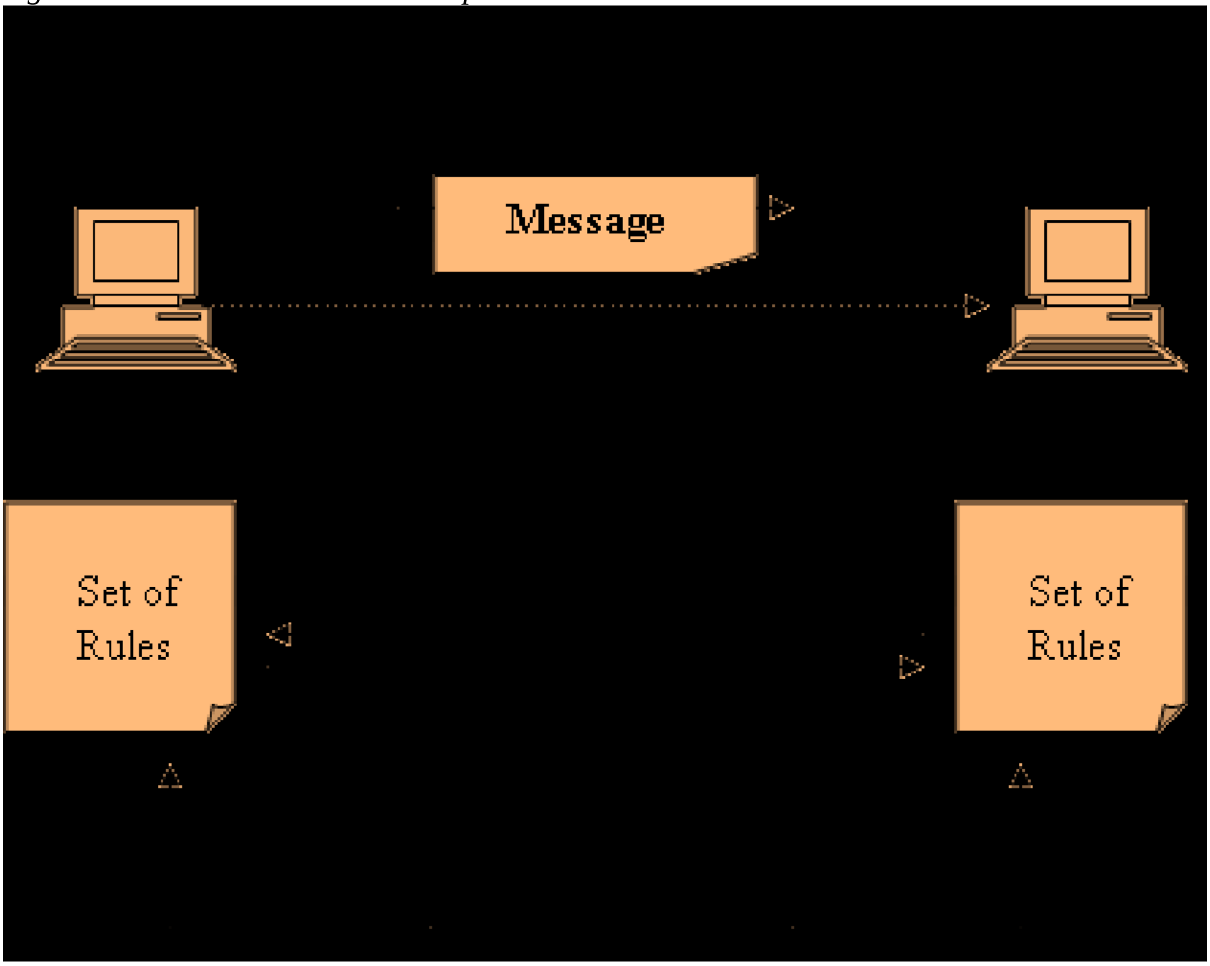
a. **Delivery:** The system must deliver data to the correct destination. Data must be received by the intended device or user and only by that device or user.
b. **Accuracy:** The system must deliver data accurately. Data that have been altered in transmission and left uncorrected are unusable. The communication systems employ error correction and recovery.
c. **Timeliness:** The system must deliver data in a timely manner. Data delivered late are useless. In the case of video, audio and voice data, timely delivery means delivering data as they are produced, in the same order that they are produced, and without significant delay. This kind of delivery is called real-time transmission.

1.2 Data Communication Component

The following are the basic components for working of a communication system and as shown in fig. 2.1.
1. Transmitter
2. Receiver
3. Medium
4. Message
5. Protocol

The transmitter sends the message and the receiver receives the message. The medium is the channel over which the message is sent and the protocol is the set of rules that guides how the data is transmitted from encoding to decoding. The message of course is central to all the components. The message is the data that is being communicated.

Fig 1.1 : Data Communication Components



Transmitter: The transmitter is the device that sends the message. It can be a computer, workstation, telephone handset, video camera, and so on.
Receiver: The receiver is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.
Medium: The transmission medium is the physical path by which a message travels from sender to receiver. It can consist of twisted pair wire, coaxial cable, fiber-optic cable, laser or radio waves (terrestrial or satellite microwave).
Message: The message is the transmission (data) to be communicated. It can consist of text, number, pictures, sound, or video or any combination of these.
Protocol: A protocol is a set of rules that governs data communication. It represents an agreement between the communicating devices. Without a protocol, two devices may be connected but not communicating, just as a person speaking German cannot be understood by a person who speaks only Japanese.

1.3. COMMUNICATION MEDIA

Following are the major communication devices which are frequently used:

a. **Wire Pairs**: Wire pairs are commonly used in local telephone communication and for short distance digital data communication. They are usually made up of copper. Using these wire pairs data transmission speed is normally 9600 bits per second in a distance of 100 metre.

b.**Twisted pair:** Twisted pair wire is the most widely used medium for telecommunication. Twisted-pair cabling consist of copper wires that are twisted into pairs. Ordinary telephone wires consist of two insulated copper wires twisted into pairs. Computer networking cabling (wired Ethernet as defined by IEEE 802.3) consists of 4 pairs of copper cabling that can be utilized for both voice and data transmission. The use of two wires twisted together helps to reduce crosstalk and electromagnetic induction. The transmission speed ranges from 2 million bits per second to 10 billion bits per second. Twisted pair cabling comes in two forms which are Unshielded Twisted Pair (UTP) and Shielded twisted-pair (STP) which are manufactured in different increments for various scenario.

c **Coaxial cable:** Coaxial cable is widely used for cable television systems, office buildings, and other work-sites for local area networks. The cables consist of copper or aluminium wire wrapped with insulating layer typically of

d.**Optical fibre:** Optical fibre cable consists of one or more filaments of glass fiber wrapped in protective layers that carries data by means of pulses of light. It transmits light which can travel over extended distances. Fiber-optic cables are not affected by electromagnetic radiation. Transmission speed may reach trillions of bits per second. The transmission speed of fiber optics is hundreds of times faster than for coaxial cables and thousands of times faster than a twisted-pair wire. This capacity may be further increased by the use of colored light, i.e., light of multiple wavelengths. Instead of carrying one message in a stream of monochromatic light impulses, this technology can carry multiple signals in a single fiber.

Wireless technologies

a. **Terrestrial microwave:** Terrestrial microwaves use Earth based transmitter and receiver. The equipment looks similar to satellite dishes. Terrestrial microwaves use low-gigahertz range, which limits all communications to line-of-sight. Path between relay stations spaced approx, 48 km (30 mi) apart. Microwave antennas are usually placed on top of buildings, towers, hills, and mountain peaks.

b.**Communications satellites:** The satellites use microwave radio signals as their telecommunications medium which are not deflected by the Earth's atmosphere. The satellites are stationed in space, typically 35,400 km (22,000 mi) (for geosynchronous satellites) above the equator. These Earth-orbiting systems are capable of receiving and relaying voice, data, and TV signals.

1.4. TYPES OF COMMUNICATION SERVICES

A term used to describe the data-handling capacity of a communication service is bandwidth. Bandwidth is the range of frequencies that is available for the transmission of data. A narrow range of frequencies in a communication system is analogous to a garden hose with a small diameter. The flow of information in such a system is restricted, just as is the flow of water in the narrow hose. Wider bandwidths permit more rapid information flow. The communication data transfer rate is measured in a unit called baud. Baud is identical to bits per second. Therefore, a rate of 300 baud is 300 bits per second. Communication companies such as American Telephone and Telegraph (AT&T) and Western Union are called common carriers, and they provide three general classes of service for both voice and data communication:

a. Narrow band handles low data volumes. Data transmission rates are from 45 to 300 baud. The low-speed devices might use narrow band communications.

b. Voice band handles moderate data transmission volumes between 300 and 9600 baud. They are used for applications ranging from operating a CRT to running a line printer. Their major application is for telephone voice communication hence, the term voice band. Broadband handles very large volumes of data. These systems provide data transmission rates of 1 million baud or more. High-speed data analysis and satellite communications are examples of broadband communication systems.

1.5. MODEM

A modem (modulator-demodulator) is a device that modulates an analog carrier signal to encode digital information, and also demodulates such a carrier signal to decode the transmitted information. The goal is to produce a signal that can be transmitted easily and decoded to reproduce the original digital data. Modems can be used over any means of transmitting analog signals, from light emitting diodes to radio. A modem modulates outgoing digital signals from a computer or other digital device to analog signals for a conventional copper twisted pair telephone line and demodulates the incoming analog signal and converts it to a digital signal for the digital device. In recent years, the 2400 bits per second modem that could carry e-mail has become obsolete. 14.4 Kbps and 28.8 Kbps modems were temporary landing places on the way to the much higher bandwidth devices and carriers of tomorrow. From early

1.6. Types of Modems

a. Landline Modems
b.Wireless Modems
c LAN Modems

a. Landline Modems:

Landline modems are modems which connect to the public switched telephone network (PSTN). To connect to PSTN, these modems have a jack known as RJ-11, or regular phone jack. A telephone cable with a RJ-11 plug connects the modem to the nearest phone jack, which also conforms to the RH-11standard. Landline modems can be further classified into the followings types:

1. Internal modems: This device is a circuit board that plugs into one of the expansion slots of the computer. Internal modems usually are cheaper than external modems, but when problems occur, fixing and troubleshooting the modem can sometimes prove to be quite difficult. The telephone line plugs into the modem port in the back of the computer. Most internal modems come installed in the computer you buy. Internal modems are more directly integrated into the computer system and, therefore, do not need any special attention. Internal modems are activated when you run a communications program and are turned off when you exit the program. This convenience is especially useful for novice users. Internal modems usually cost less than external modems, but the price difference is usually small. The major disadvantage with internal modems is their location: inside the computer. When you want to replace an internal modem you have to go inside the computer case to make the switch.

2. External modems: This device is attached to the back of the computer by way of a cable that plugs into the modem port. It is usually less expensive and very portable. It can be used with other computers very easily by unplugging it and plugging it into another computer. This is the simplest type of modem to install because you don't have to open the computer. External modems have their own power supply and connect with a cable to a computer's serial port. The telephone line plugs into a socket on the rear panel of the modem. Because external modems have their own power supply, you can turn off the modem to break an online connection quickly without powering down the computer. Another advantage over an internal modem is that an external modem's separate power supply does not drain any power from the computer. You also can monitor your modem's connection activity by watching the status lights.

3. Voice/data/fax modems: This device can be hooked up to your telephone and used to send information to your computer. Your computer can also send information to a fax machine. Most computer modems are modems with faxing capabilities.

4. PC Card modem: These modems, designed for portable computers, are the size of a credit card and fit into the PC Card slot on notebook and handheld computers. These modems are removed when the modem is not needed. Except for their size, PC Card modems are like a combination of external and internal modems. These devices are plugged directly into an external slot in the portable computer, so no cable is required other than the telephone line connection. The cards are powered by the computer, which is fine unless the computer is battery-operated. Running a PC Card modem while the portable computer is operating on battery power drastically decreases the life of your batteries.

b. Wireless Modems:

Wireless modems are radio transmitters/receivers installed into mobile computing devices (i.e. devices that are used while you are moving such as mobile phones, laptops etc.) Using wireless modems, one can connect to a network while being mobile. Unlike landline modems, wireless modems do not plug into an RJ-11 jack.

C. LAN Modems:

LAN modems allow shared remote access to LAN (Local Area Network) resources. LAN modem comes fully preconfigured for single particular network architecture such as Ethernet or Token Ring and/or particular network software such as IPX, NetBIOS, NetBEUI etc.

1.7. Working of Modem

Modems convert analog data transmitted over phone lines into digital data so that computers can read; they also convert digital data into analog data so it can be transmitted. This process involves modulating and demodulating the computer's digital signals into analog signals that travel over the telephone lines. In other words, the modem translates computer data into the language used by telephones and then reverses the process to translate the responding data back into computer language.

2.0 COMPUTER NETWORK

A computer network is interconnection of various computer systems located at different places. In computer network two or more computers are linked together with a medium and data communication devices for the purpose of communicating data and sharing resources. The computer that provides resources to other computers on a network is known as server. In the network the individual computers, which access shared network resources, are known as nodes.

2.1. Characteristics of a Computer Network

The primary purpose of a computer network is to share resources:

a. You can play a CD music from one computer while sitting on another computer

b. You may have a computer that doesn't have a DVD or BluRay (BD) player. In this case, you can place a movie disc (DVD or BD) on the computer that has the player, and then view the movie on a computer that lacks the player

c. You may have a computer with a CD/DVD/BD writer or a backup system but the other computer doesn't have it. In this case, you can burn discs or make backups on a computer that has one of these but using data from a computer that doesn't have a disc writer or a backup system

d.You can connect a printer (or a scanner, or a fax machine) to one computer and let other computers of the network print (or scan, or fax) to that printer (or scanner, or fax machine)

e. You can place a disc with pictures on one computer and let other computers access those pictures

f. You can create files and store them in one computer, then access those files from the other computer(s) connected to it.

2.3. Concept of Networking:

A computer network, often simply referred to as a network, is a collection of hardware components and computers interconnected by communication channels that allow sharing of resources and information. Networks may be classified according to a wide variety of characteristics such as the medium used to transport the data, communications protocol used, scale, topology, and organizational scope. The rules and data formats for exchanging information in a computer network are defined by communications protocols.

2.4. Properties of Network

1. Facilitate communications:

Using a network, people can communicate efficiently and easily via email, instant messaging, chat rooms, telephone, video telephone calls, and video conferencing.

2. Permit sharing of files, data, and other types of information

In a network environment, authorized users may access data and information stored on other computers on the network. The capability of providing access to data and information on shared storage devices is an important feature of many networks.

3. Share network and computing resources

In a networked environment, each computer on a network may access and use resources provided by devices on the network, such as printing a document on a shared network printer. Distributed computing uses computing resources across a network to accomplish tasks.

2.5. BENEFITS OF NETWORK

a. **File sharing**: Network file sharing between computers gives you more flexibility than using floppy drives or Zip drives. Not only can you share photos, music files, and documents, you can also use a home network to save copies of all of your important data on a different computer. Backups are one of the most critical yet overlooked tasks in home networking.

b. **Printer / peripheral sharing:** Once a home network is in place, it's then easy to set up all of the computers to share a single printer. No longer will you need to bounce from one system or another just to print out an email message. Other computer peripherals can be shared similarly such as network scanners, Web cams, and CD burners.

c. **Internet connection sharing:** Using a home network, multiple family members can access the Internet simultaneously without having to pay an ISP for multiple accounts. You will notice the Internet connection slows down when several people share it, but broadband Internet can handle the extra load with little trouble.

d. **Multi-player games**: Many popular home computer games support LAN mode where friends and family can

play together, if they have their computers networked.
e. **Internet telephone service**: Voice over IP (VoIP) services allows you to make and receive phone calls through your home network across the Internet.
e. **Home entertainment:** Newer home entertainment products such as digital video recorders (DVRs) and video game consoles now support either wired or wireless home networking. Having these products integrated into your network enables online Internet gaming, video sharing and other advanced features.

2.6. TYPES OF NETWORK

There are many different types of networks. However, from an end user's point of view there are three basic types:
a Local Area Network
b Wide Area Network
c Metropolitan Area Network

2.6.1 Local Area Network (LAN):

A local area network (LAN) supplies networking capability to a group of computers in close proximity to each other such as in an office building, a school, or a home. A LAN is useful for sharing resources like files, printers, games or other applications. A LAN in turn often connects to other LANs, and to the Internet or other WAN. Most local area networks are built with relatively inexpensive hardware such as Ethernet cables, network adapters, and hubs.
Wireless LAN and other more advanced LAN hardware options also exist. LAN is a computer network that spans a relatively small area. Most LANs are confined to a single building or group of buildings. However, one LAN can be connected to other LANs over any distance via telephone lines and radio waves. Most LANs (as shown in Fig. 2.2) connect workstations and personal computers. Each node (individual computer) in a LAN has its own CPU with which it executes programs, but it is also able to access data and devices anywhere on the LAN. This means that many users can share expensive devices, such as laser printers, as well as data. Users can also use the LAN to communicate with each other, by sending e-mail or engaging in chat sessions. There are many different types of LANs-token ring networks, Ethernets, and ARC nets being the most common for PCs.

2.6.2 Wide Area Networks (WANs)

The term Wide Area Network (WAN) usually refers to a network which covers a large geographical area, and use communications circuits to connect the intermediate nodes. A major factor impacting WAN design and performance is a requirement that they lease communications circuits from telephone companies or other communications carriers. Transmission rates are typically 2 Mbps, 34 Mbps, 45 Mbps, 155 Mbps, 625 Mbps (or sometimes considerably more).

2.6.3 Metropolitan Area Network (MAN)

A Metropolitan Area Network (MAN) is one of a number of types of networks (see also LAN and WAN). A MAN is a relatively new class of network, it serves a role similar to an ISP, but for corporate users with large LANs.

2.7 IMPORTANT TERMS USED IN NETWORKING

a. Voice Messaging

It is a new communication approach which is similar to electronic mail except that it is audio message rather than text messages that are processed. A sender speaks into a telephone rather than typing, giving the name of the recipient and the message. That sender's voice signal is then digitized and stored.

The system can then either deliver the message at a specified time in future or it can be retrieved from a database by the recipient. The message is reconverted back into its analog format when it is delivered or retrieved so that the recipient hears it as the original sender's voice on a telephone. Voice messaging requires a computer with an ability to store the audio messages in digital form and then convert them back in an audio form upon verification. Each user has a voice mailbox in secondary storage and special equipment converts the audio message to and from the digital form. The main advantage of voice mail over electronic mail is that the sender does not have to type. Voice mail also makes it easy to include people in the firm's environment in a communication network.

b. Hub

A hub is typically the least expensive, least intelligent, and least complicated. Its job is very simple: anything that comes in one port is sent out to the others. Every computer connected to the hub "sees" everything that every other computer on the hub sees. The hub itself is blissfully ignorant of the data being transmitted. For years, simple hubs have been quick and easy ways to connect computers in small networks.

c. Switch

A switch does essentially what a hub does but more efficiently. By paying attention to the traffic that comes across it, it can "learn" where particular addresses are. For example, if it sees traffic from machine A coming in on port 2, it now knows that machine A is connected to that port and that traffic to machine A needs to only be sent to that port and not any of the others. The net result of using a switch over a hub is that most of the network traffic only goes where it needs to rather than to every port. On busy networks this can make the network significantly faster.

d. Router

A router is the smartest and most complicated of the bunch. Routers come in all shapes and sizes from the small four-port broadband routers that are very popular right now to the large industrial strength devices that drive the internet itself. A simple way to think of a router is as a computer that can be programmed to understand, possibly manipulate, and route the data its being asked to handle. For example, broadband routers include the ability to "hide" computers behind a type of firewall which involves slightly modifying the packets of network traffic as they traverse the device. All routers include some kind of user interface for configuring how the router will treat traffic. The really large routers include the equivalent of a full-blown programming language to describe how they should operate as well as the ability to communicate with other routers to describe or determine the best way to get network traffic from point A to point B.

d. Network Repeater

A repeater connects two segments of your network cable. It retimes and regenerates the signals to proper amplitudes and sends them to the other segments. When talking about, Ethernet topology, you are probably talking about using a hub as a repeater. Repeaters require a small amount of time to regenerate the signal. This can cause a propagation delay which can affect network communication when there are several repeaters in a row. Many network architectures limit the number of repeaters that can be used in a row. Repeaters work only at the physical layer of the OSI network model.

e. Bridge

A bridge reads the outermost section of data on the data packet, to tell where the message is going. It reduces the traffic on other network segments, since it does not send all packets. Bridges can be programmed to reject packets from particular networks. Bridging occurs at the data link layer of the OSI model, which means the bridge cannot read IP addresses, but only the outermost hardware address of the packet. In our case the bridge can read the Ethernet data which gives the hardware address of the destination address, not the IP address. Bridges forward all broadcast messages. Only a special bridge called a translation bridge will allow two networks of different architectures to be connected. Bridges do not normally allow connection of networks with different architectures. The hardware address is also called the MAC (media access control) address. To determine the network segment a MAC address belongs to, bridges use one of the following:

f. Transparent Bridging: They build a table of addresses (bridging table) as they receive packets. If the address is not in the bridging table, the packet is forwarded to all segments other than the one it came from. This type of bridge is used on Ethernet networks.

g. Source route bridging: The source computer provides path information inside the packet. This is used on Token Ring networks

h. Router

There is a device called a router which will function similar to a bridge for network transport protocols that are not routable, and will function as a router for routable protocols. It functions at the network and data link layers of the OSI network model.

i. Gateway

A gateway can translate information between different network data formats or network architectures. It can translate TCP/IP to AppleTalk so computers supporting TCP/IP can communicate with Apple brand computers. Most gateways operate at the application layer, but can operate at the network or session layer of the OSI model. Gateways will start at the lower level and strip information until it gets to the required level and repackage the information and work its way back toward the hardware layer of the OSI model.

2.8. TELECONFERENCING

The term teleconferencing refers to electronic meetings that involve people who are at physically different sites. Telecommunication technology system allows meeting participants to interact with one another without travelling to the same location. Three different types of teleconferencing exist: audio teleconferencing, video teleconferencing and computer conferencing.

2.8.1. Audio Conferencing

Audio conferencing is the use of voice communication equipments to establish an audio link between geographically dispersed persons, one that allows them to conduct a conference. The conference call was the first form of audit conferencing and is still in use. Some firms install more elaborate systems consisting of private, high-quality audio communications circuits that can be activated with the flip of a switch. Audio conferencing does not require a computer. It only requires a two-way audio communications facility. Audio conferencing is best suited for firms that are spread over a wide area. However, since it is a form of synchronous communication that requires all participants to be present at the same time, it is difficult to schedule conferences when time zones are far apart.

2.8.2. Video Conferencing:
Video conferencing is the use of television equipment to link geographically dispersed conference participants. The equipment provides both sound and picture. Like audio conferencing, video conferencing also does not necessarily require a computer. With video conferencing, participants can see and hear each other. Generally, participants gather in relatively expensive, specially equipped rooms that can handle the complexities of simultaneous video and audio transmission. There are three possible video conferencing configurations.

2.8.3. One-Way Video and Audio: Video and audio signals are sent from a single transmitting site to one or more receiving sites. This is a good way for a project leader to disseminate information to team members at remote locations.

2.8.4. One-Way Video and Two-Way Audio: People at the receiving sites can talk to people at the transmitting site, while everyone views the same video images.

2.8.5. Two-Way Video and Audio: The video and audio communications between all sites are two-way. Although this is the most effective of the electronically aided conferencing approaches, it can be the most expensive as well.

2.8.6. Computer Conferencing
A third form of electronic conferencing is computer conferencing. There is a fine line between this system and Email. Both use the same software and hardware. Two factors determine this application, who uses the system, and the subject matter. E-mail is available to any one who has access to the network and that includes practically everyone in the office. Also, the E-mail system can be used for any purpose. Computer conferencing, on the other hand, is the use of a networked computer that allows particular task. Computer conferencing is more disciplined form of E-mail. Unlike an audio conference, a, computer conference group can consist of large number of participants. One of the largest computer conferences was formed within IBM to include anyone who had an interest in the IBM PC. Its members exceeded 40,000, and there were over 4,000 separate topic areas. Computer conferencing differs from audio and video conferencing because it can be used within a single geographic site. A person can use computer conferencing to communicate with someone in the office next door.

3.0. Network planning and design

Network planning and design is an iterative process, encompassing topological design, network-synthesis, and network-realization, and is aimed at ensuring that a new telecommunications network or service meets the needs of the subscriber and operator. The process can be tailored according to each new network or service.

3.1. A network planning methodology

A traditional network planning methodology in the context of business decisions involves five layers of planning, namely:

- need assessment and resource assessment
- short-term network planning
- IT resource
- long-term and medium-term network planning
- operations and maintenance.[1]

Each of these layers incorporates plans for different time horizons, i.e. the business planning layer determines the planning that the operator must perform to ensure that the network will perform as required for its intended life-span. The Operations and Maintenance layer, however, examines how the network will run on a day-to-day basis.

The network planning process begins with the acquisition of external information. This includes:

- forecasts of how the new network/service will operate;
- the economic information concerning costs; and
- the technical details of the network's capabilities.

Planning a new network/service involves implementing the new system across the first four layers of the OSI Reference Model. Choices must be made for the protocols and transmission technologies.

Network planning process involves three main steps:

- **Topological design**: This stage involves determining where to place the components and how to connect them. The (topological) optimisation methods that can be used in this stage come from an area of mathematics called Graph Theory. These methods involve determining the costs of transmission and the cost of switching, and thereby determining the optimum connection matrix and location of switches and concentrators.[1]
- **Network-synthesis**: This stage involves determining the size of the components used, subject to performance criteria such as the Grade of Service (GOS). The method used is known as "Nonlinear Optimisation", and involves determining the topology, required GoS, cost of transmission, etc., and using this information to calculate a routing plan, and the size of the components.[1]
- **Network realization**: This stage involves determining how to meet capacity requirements, and ensure reliability within the network. The method used is known as

"Multi-commodity Flow Optimisation", and involves determining all information relating to demand, costs and reliability, and then using this information to calculate an actual physical circuit plan.

These steps are performed iteratively in parallel with one another.

The role of forecasting

During the process of Network Planning and Design, estimates are made of the expected traffic intensity and traffic load that the network must support.[1] If a network of a similar nature already exists, traffic measurements of such a network can be used to calculate the exact traffic load.[2] If there are no similar networks, then the network planner must use telecommunications forecasting methods to estimate the expected traffic intensity.[1]

The forecasting process involves several steps:[1]

- Definition of problem;
- Data acquisition;
- Choice of forecasting method;
- Analysis/Forecasting;
- Documentation and analysis of results.

Dimensioning

Dimensioning a new network determines the minimum capacity requirements that will still allow the Teletraffic Grade of Service (GoS) requirements to be met.[1][2] To do this, dimensioning involves planning for peak-hour traffic, i.e. that hour during the day during which traffic intensity is at its peak.[1]

The dimensioning process involves determining the network's topology, routing plan, traffic matrix, and GoS requirements, and using this information to determine the maximum call handling capacity of the switches, and the maximum number of channels required between the switches.[1] This process requires a complex model that simulates the behavior of the network equipment and routing protocols.

A dimensioning rule is that the planner must ensure that the traffic load should never approach a load of 100 percent.[1] To calculate the correct dimensioning to comply with the above rule, the planner must take on-going measurements of the network's traffic, and continuously maintain and upgrade resources to meet the changing requirements.[1][2] Another reason for **over-provisioning** is to make sure that traffic can be rerouted in case a failure occurs in the network.

Because of the complexity of network dimensioning, this is typically done using specialized software tools. Whereas researchers typically develop custom software to study a particular problem, network operators typically make use of commercial network planning software.

Traffic engineering

Compared to network engineering, which adds resources such as links, routers and switches into the network, traffic engineering targets changing traffic paths on the existing network to alleviate traffic congestion or accommodate more traffic demand.

This technology is critical when the cost of network expansion is prohibitively high and network load is not optimally balanced. The first part provides financial motivation for traffic engineering while the second part grants the possibility of deploying this technology.

Survivability

Network survivability enables the network to maintain maximum network connectivity and quality of service under failure conditions. It has been one of the critical requirements in network planning and design. It involves design requirements on topology, protocol, bandwidth allocation, etc.. Topology requirement can be maintaining a minimum two-connected network against any failure of a single link or node. Protocol requirements include using dynamic routing protocol to reroute traffic against network dynamics during the transition of network dimensioning or equipment failures. Bandwidth allocation requirements pro-actively allocate extra bandwidth to avoid traffic loss under failure conditions. This topic has been actively studied in conferences, such as the International Workshop on Design of Reliable Communication Networks.[3]

Tools

There are a wide variety of tools available for network planning and design depending on the technologies being used. These include:

- OPNET
- NetSim

4.0. Serial communication

In telecommunication and data transmission, **serial communication** is the process of sending data one bit at a time, sequentially, over a communication channel or computer bus. This is in contrast to parallel communication, where several bits are sent as a whole, on a link with several parallel channels.

Serial communication is used for all long-haul communication and most computer networks, where the cost of cable and synchronization difficulties make parallel communication impractical. Serial computer buses are becoming more common even at shorter distances, as improved signal integrity and transmission speeds in newer serial technologies have begun to outweigh the parallel bus's advantage of simplicity (no need for serializer and deserializer, or SerDes) and to outstrip its disadvantages (clock skew, interconnect density). The migration from PCI to PCI Express is an example.

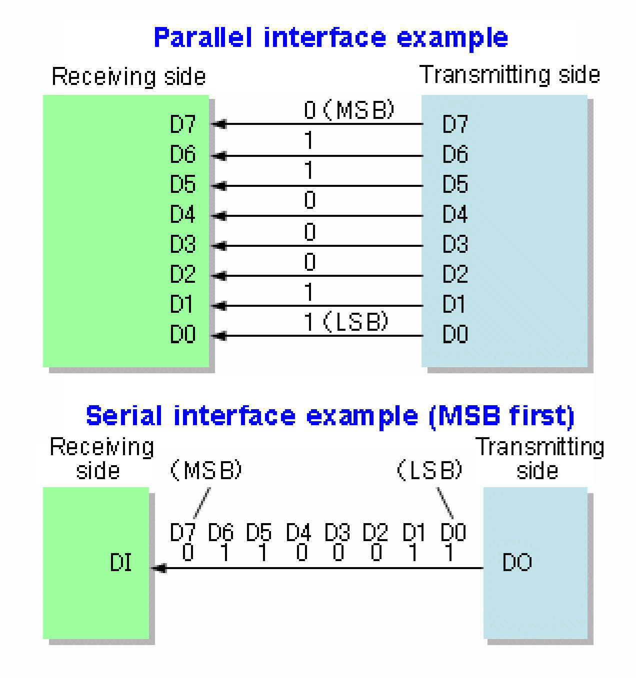


Fig 4.1. Parallel versus serial communication.

Cables

Many serial communication systems were originally designed to transfer data over relatively large distances through some sort of data cable.

Practically all long-distance communication transmits data one bit at a time, rather than in parallel, because it reduces the cost of the cable. The cables that carry this data (other than "the" serial cable) and the computer ports they plug into are usually referred to with a more specific name, to reduce confusion.

Keyboard and mouse cables and ports are almost invariably serial—such as PS/2 port and Apple Desktop Bus and USB.

The cables that carry digital video are almost invariably serial—such as coax cable plugged into a HD-SDI port, a webcam plugged into a USB port or Firewire port, Ethernet cable connecting an IP camera to a Power over Ethernet port, FPD-Link, etc.

Other such cables and ports, transmitting data one bit at a time, include Serial ATA, Serial SCSI, Ethernet cable plugged into Ethernet ports, the Display Data Channel using previously reserved pins of the VGA connector or the DVI port or the HDMI port.

Serial buses

Many communication systems were generally designed to connect two integrated circuits on the same printed circuit board, connected by signal traces on that board (rather than external cables).

Integrated circuits are more expensive when they have more pins. To reduce the number of pins in a package, many ICs use a serial bus to transfer data when speed is not important. Some examples of such low-cost serial buses include SPI, I²C, DC-BUS, UNI/O, and 1-Wire.

Serial versus parallel

The communication links, across which computers (or parts of computers) talk to one another, may be either serial or parallel. A parallel link transmits several streams of data simultaneously along multiple channels (e.g., wires, printed circuit tracks, or optical fibres); whereas, a serial link transmits only a single stream of data.

Although a serial link may seem inferior to a parallel one, since it can transmit less data per clock cycle, it is often the case that serial links can be clocked considerably faster than parallel links in order to achieve a higher data rate. Several factors allow serial to be clocked at a higher rate:

- Clock skew between different channels is not an issue (for unclocked asynchronous serial communication links).
- A serial connection requires fewer interconnecting cables (e.g., wires/fibres) and hence occupies less space. The extra space allows for better isolation of the channel from its surroundings.
- Crosstalk is less of an issue, because there are fewer conductors in proximity.

In many cases, serial is cheaper to implement than parallel. Many ICs have serial interfaces, as opposed to parallel ones, so that they have fewer pins and are therefore less expensive.

4.1. Comparison of synchronous and asynchronous signaling

Synchronous and **asynchronous transmissions** are two different methods of transmission synchronization. Synchronous transmissions are synchronized by an external clock, while asynchronous transmissions are synchronized by special signals along the transmission medium.

The need for synchronization

Whenever an electronic device transmits digital (and sometimes analog) data to another, there must be a certain rhythm established between the two devices, i.e., the receiving device must have some way of, within the context of the fluctuating signal that it's receiving, determining where each unit of data begins and where it ends.

Methods of synchronization

There are two ways to synchronize the two ends of the communication.

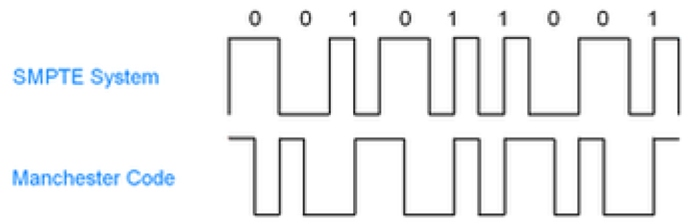
The synchronous signalling methods use two different signals. A pulse on one signal indicates when another bit of information is ready on the other signal.

The asynchronous signalling methods use only one signal. The receiver uses transitions on that signal to figure out the transmitter bit rate ("autobaud") and timing, and set a local clock to the proper timing, typically using a phase-locked loop (PLL) to synchronize with the transmission rate. A pulse from the local clock indicates when another bit is ready.

Synchronous transmission

In synchronous communications, the stream of data to be transferred is encoded as fluctuating voltage levels in one wire (the 'DATA'), and a periodic pulse of voltage on a separate wire (called the "CLOCK" or "STROBE") which tells the receiver "the current DATA bit is 'valid' at this moment in time".

Practically all parallel communications protocols use synchronous transmission. For example, in a computer, address information is transmitted synchronously—the address bits over the address bus, and the read or write 'strobe's of the control bus.



Single-wire synchronous signalling

A logical one is indicated when there are two transitions in the same time frame as a zero. In the Manchester coding a transition from low to high indicates a one and a transition from high to low indicates a zero. When there are successive ones or zeros, an opposite transition is required on the edge of the time frame to prepare for the next transition and signal.

Asynchronous transmission

The most common asynchronous signalling, asynchronous start-stop signalling, uses a near-constant 'bit' timing (+/- 5% local oscillator required at both end of the connection). Using this, the receiver detects the 'first' edge transition... (the START BIT), then waits 'half a bit duration' then reads A further delay of one 'whole bit duration' is executed before the next data bit is 'read' - repeating for length of the whole serial word (typically 7/8-data bits). Finally an optional STOP bit is appended to identify the end of the data word.

The word structure used in typical asynchronous serial communications is START-DATA[0:7]-STOP[0:1] (followed by an optional PARITY bit). These formatting variables are specified when configuring the transmit and receive nodes before communications take place. The bit duration is determined from the nominated 'bit rate' in bps... 300, 1200, 9600, 19200, 115200 etc. The use of the word BAUD is not strictly correct in the modern application of serial channels.

Special level & timing conditions are detected to identify an open-circuit condition (BREAK)

The sync token might be a single pulse (a "start bit" as noted above), or it may be a more complicated syncword or self-synchronizing code such as HDLC or 8B/10B encoding.

4.2. Advantages and disadvantages

	Advantages	**Disadvantages**
Asynchronous transmission	• Simple, doesn't require synchronization of both communication sides • Cheap, because asynchronous transmission requires less hardware • Setup is faster than other transmissions, so well suited for applications where messages are generated at irregular intervals, for example data entry from the keyboard, and the speed depends on different applications.	• Large relative overhead, a high proportion of the transmitted bits are uniquely for control purposes and thus carry no useful information
Synchronous transmission	• Lower overhead and thus, greater throughput	• Slightly more complex • Hardware is more expensive

5.0. Fourier analysis

In mathematics, **Fourier analysis** is the study of the way general functions may be represented or approximated by sums of simpler trigonometric functions. Fourier analysis grew from the study of Fourier series, and is named after Joseph Fourier, who showed that representing a function as a sum of trigonometric functions greatly simplifies the study of heat transfer.

Today, the subject of Fourier analysis encompasses a vast spectrum of mathematics. In the sciences and engineering, the process of decomposing a function into oscillatory components is often called Fourier analysis, while the operation of rebuilding the function from these pieces is known as **Fourier synthesis**. For example, determining what component frequencies are present in a musical note would involve computing the Fourier transform of a sampled musical note. One could then re-synthesize the same sound by including the frequency components as revealed in the Fourier analysis. In mathematics, the term *Fourier analysis* often refers to the study of both operations.

The decomposition process itself is called a Fourier transformation. Its output, the Fourier transform, is often given a more specific name, which depends on the domain and other properties of the function being transformed. Moreover, the original concept of Fourier analysis has been extended over time to apply to more and more abstract and general situations, and the general field is often known as harmonic analysis. Each transform used for analysis (see list of Fourier-related transforms) has a corresponding inverse transform that can be used for synthesis.

5.1. Applications of Fourier analysis

Fourier analysis has many scientific applications – in physics, partial differential equations, number theory, combinatorics, signal processing, digital image processing, probability theory, statistics, forensics, option pricing, cryptography, numerical analysis, acoustics, oceanography, sonar, optics, diffraction, geometry, protein structure analysis, and other areas.

This wide applicability stems from many useful properties of the transforms:

- The transforms are linear operators and, with proper normalization, are unitary as well (a property known as Parseval's theorem or, more generally, as the Plancherel theorem, and most generally via Pontryagin duality) (Rudin 1990).
- The transforms are usually invertible.
- The exponential functions are eigen functions of differentiation, which means that this representation transforms linear differential equations with constant coefficients into ordinary algebraic ones (Evans 1998). Therefore, the behavior of a linear time-invariant system can be analyzed at each frequency independently.
- By the convolution theorem, Fourier transforms turn the complicated convolution operation into simple multiplication, which means that they provide an efficient way to compute convolution-based operations such as polynomial multiplication and multiplying large numbers (Knuth 1997).
- The discrete version of the Fourier transform (see below) can be evaluated quickly on computers using Fast Fourier Transform (FFT) algorithms. (Conte & de Boor 1980)

In forensics, laboratory infrared spectrophotometers use Fourier transform analysis for measuring the wavelengths of light at which a material will absorb in the infrared spectrum. The FT method is used to decode the measured signals and record the wavelength data. And by using a computer, these Fourier calculations are rapidly carried out, so that in a matter of seconds, a computer-operated FT-IR instrument can produce an infrared absorption pattern comparable to that of a prism instrument.

Fourier transformation is also useful as a compact representation of a signal. For example, JPEG compression uses a variant of the Fourier transformation (discrete cosine transform) of small square pieces of a digital image. The Fourier components of each square are rounded to lower arithmetic precision, and weak components are eliminated entirely, so that the remaining components can be stored very compactly. In image reconstruction, each image square is reassembled from the preserved approximate Fourier-transformed components, which are then inverse-transformed to produce an approximation of the original image.

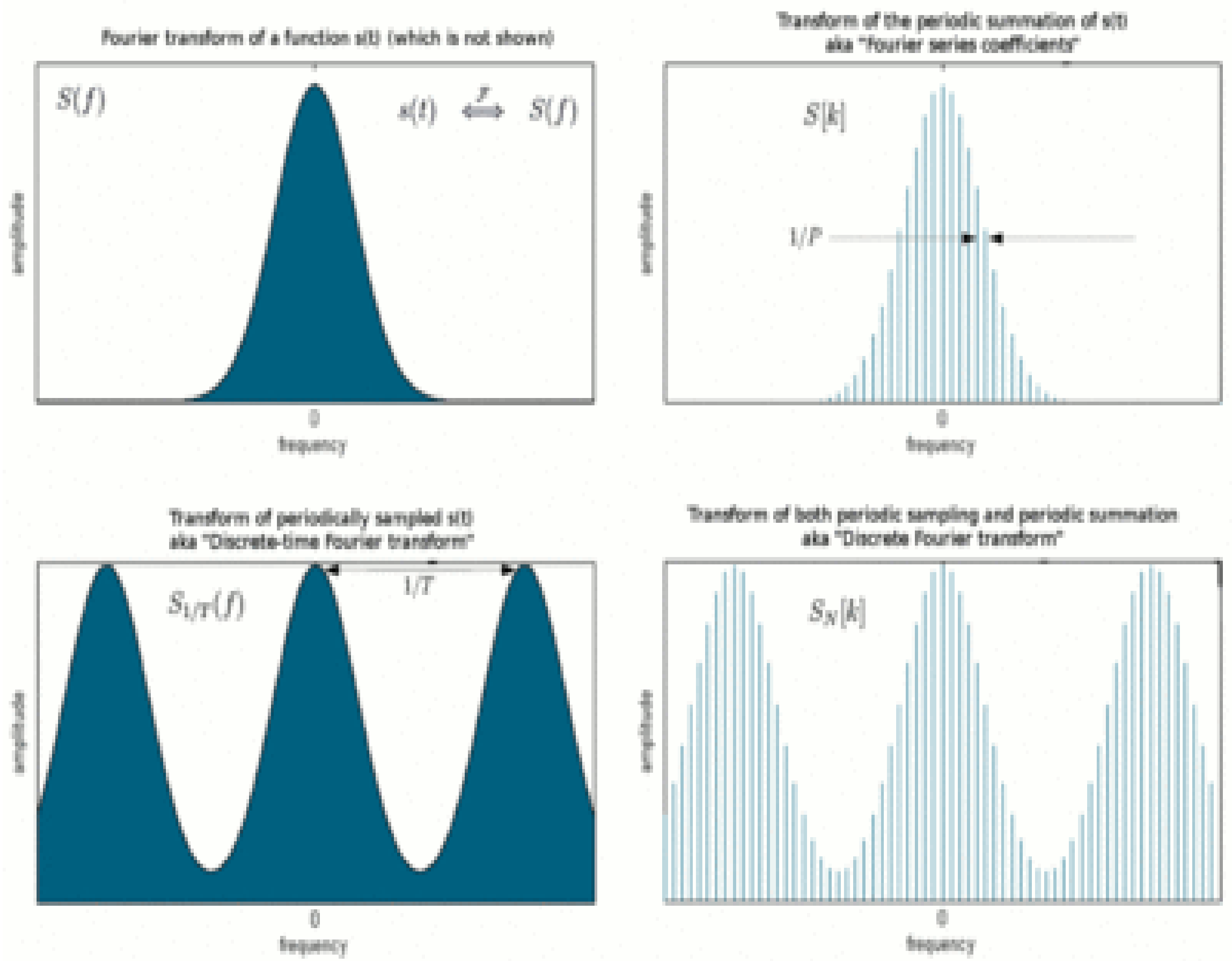
5.2. Applications in signal processing

When processing signals, such as audio, radio waves, light waves, seismic waves, and even images, Fourier analysis can isolate narrowband components of a compound waveform, concentrating them for easier detection or removal. A large family of signal processing techniques consist of Fourier-transforming a signal, manipulating the Fourier-transformed data in a simple way, and reversing the transformation.[3]

Some examples include:

- Equalization of audio recordings with a series of bandpass filters;
- Digital radio reception without a superheterodyne circuit, as in a modern cell phone or radio scanner;
- Image processing to remove periodic or anisotropic artifacts such as jaggies from interlaced video, strip artifacts from strip aerial photography, or wave patterns from radio frequency interference in a digital camera;
- Cross correlation of similar images for co-alignment;
- X-ray crystallography to reconstruct a crystal structure from its diffraction pattern;
- Fourier transform ion cyclotron resonance mass spectrometry to determine the mass of ions from the frequency of cyclotron motion in a magnetic field;
- Many other forms of spectroscopy, including infrared and nuclear magnetic resonance spectroscopies;
- Generation of sound spectrograms used to analyze sounds;
- Passive sonar used to classify targets based on machinery noise.

Variants of Fourier analysis



A Fourier transform and 3 variations caused by periodic sampling (at interval T) and/or periodic summation (at interval P) of the underlying time-domain function. The relative computational ease of the DFT sequence and the insight it gives into $S(f)$ make it a popular analysis tool.

(Continuous) Fourier transform
Main article: Fourier transform

Most often, the unqualified term Fourier transform refers to the transform of functions of a continuous real argument, and it produces a continuous function of frequency, known as a *frequency distribution*. One function is transformed into another, and the operation is reversible. When the domain of the input (initial) function is time (t), and the domain of the output (final) function is ordinary frequency, the transform of function $s(t)$ at frequency f is given by the complex number:

$$S(f) = \int_{-\infty}^{\infty} s(t) \cdot e^{-2i\pi ft}\, dt.$$

Evaluating this quantity for all values of f produces the *frequency-domain* function. Then $s(t)$ can be represented as a recombination of complex exponentials of all possible frequencies:

$$s(t) = \int_{-\infty}^{\infty} S(f) \cdot e^{2i\pi ft}\, df,$$

which is the inverse transform formula. The complex number, $S(f)$, conveys both amplitude and phase of frequency f.

See Fourier transform for much more information, including:

- conventions for amplitude normalization and frequency scaling/units
- transform properties
- tabulated transforms of specific functions
- an extension/generalization for functions of multiple dimensions, such as images.

Fourier series

Main article: Fourier series

The Fourier transform of a periodic function, $s_P(t)$, with period P, becomes a Dirac comb function, modulated by a sequence of complex coefficients:

$$S[k] = \frac{1}{P} \int_P s_P(t) \cdot e^{-2i\pi \frac{k}{P} t}\, dt$$

for all integer values of k, and where $\int_P$ is the integral over any interval of length P.

The inverse transform, known as Fourier series, is a representation of $s_P(t)$ in terms of a summation of a potentially infinite number of harmonically related sinusoids or complex exponential functions, each with an amplitude and phase specified by one of the coefficients:

$$s_P(t) = \sum_{k=-\infty}^{\infty} S[k] \cdot e^{2i\pi \frac{k}{P} t} \quad \stackrel{\mathcal{F}}{\Longleftrightarrow} \quad \sum_{k=-\infty}^{+\infty} S[k]\, \delta\left(f - \frac{k}{P}\right).$$

When $s_P(t)$, is expressed as a periodic summation of another function, $s(t)$:

$$s_P(t) \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} s(t - mP),$$

the coefficients are proportional to samples of $S(f)$ at discrete intervals of $1/\boxed{P}$

$$S[k] = \frac{1}{P} \cdot S\left(\frac{k}{P}\right).$$

A sufficient condition for recovering $s(t)$ (and therefore $S(f)$) from just these samples (i.e. from the Fourier series) is that the non-zero portion of $s(t)$ be confined to a known interval of duration P, which is the frequency domain dual of the Nyquist–Shannon sampling theorem.

See Fourier series for more information, including the historical development.

Discrete-time Fourier transform (DTFT)

The DTFT is the mathematical dual of the time-domain Fourier series. Thus, a convergent periodic summation in the frequency domain can be represented by a Fourier series, whose coefficients are samples of a related continuous time function:

$$S_{\frac{1}{T}}(f) \stackrel{\text{def}}{=} \underbrace{\sum_{k=-\infty}^{\infty} S\left(f - \frac{k}{T}\right) \equiv \overbrace{\sum_{n=-\infty}^{\infty} s[n] \cdot e^{-2i\pi f n T}}^{\text{Fourier series (DTFT)}}}_{\text{Poisson summation formula}} = \mathcal{F}\left\{\sum_{n=-\infty}^{\infty} s[n]\, \delta(t-nT)\right\},$$

which is known as the DTFT. Thus the DTFT of the $s[n]$ sequence is also the Fourier transform of the modulated Dirac comb function.[note 2]

The Fourier series coefficients (and inverse transform), are defined by:

$$s[n] \stackrel{\text{def}}{=} T \int_{\frac{1}{T}} S_{\frac{1}{T}}(f) \cdot e^{2i\pi f n T}\, df = T \underbrace{\int_{-\infty}^{\infty} S(f) \cdot e^{2i\pi f n T}\, df}_{\stackrel{\text{def}}{=} s(nT)}.$$

Parameter T corresponds to the sampling interval, and this Fourier series can now be recognized as a form of the Poisson summation formula. Thus we have the important result that when a discrete data sequence, $s[n]$, is proportional to samples of an underlying continuous function, $s(t)$, one can observe a periodic summation of the

continuous Fourier transform, $S(f)$. That is a cornerstone in the foundation of digital signal processing. Furthermore, under certain idealized conditions one can theoretically recover $S(f)$ and $s(t)$ exactly. A sufficient condition for perfect recovery is that the non-zero portion of $S(f)$ be confined to a known frequency interval of width $1/T$. When that interval is $[-1/2T, 1/2T]$, the applicable reconstruction formula is the Whittaker–Shannon interpolation formula.

Another reason to be interested in $S_{1/T}(f)$ is that it often provides insight into the amount of aliasing caused by the sampling process.

Applications of the DTFT are not limited to sampled functions. See Discrete-time Fourier transform for more information on this and other topics, including:

- normalized frequency units

- windowing (finite-length sequences)
- transform properties
- tabulated transforms of specific functions

Discrete Fourier transform (DFT)
Main article: Discrete Fourier transform

Similar to a Fourier series, the DTFT of a periodic sequence, $s_N[n]$, with period N, becomes a Dirac comb function, modulated by a sequence of complex coefficients (see DTFT/Periodic data):

$$S[k] = \sum_N s_N[n] \cdot e^{-2i\pi \frac{k}{N} n},$$

where $\sum_N$ is the sum over any n-sequence of length N.

The $S[k]$ sequence is what is customarily known as the DFT of s_N. It is also N-periodic, so it is never necessary to compute more than N coefficients. The inverse transform is given by:

$$s_N[n] = \frac{1}{N} \sum_N S[k] \cdot e^{2i\pi \frac{n}{N} k},$$

where $\sum_N$ is the sum over any k-sequence of length N.

When $s_N[n]$ is expressed as a periodic summation of another function:

$$s_N[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} s[n-mN],$$

and $s[n] \stackrel{\text{def}}{=} s(nT),$

the coefficients are proportional to samples of $S_{1/T}(f)$ at discrete intervals of $1/P = 1/NT$:

$$S[k] = \frac{1}{T} \cdot S_{\frac{1}{T}}\left(\frac{k}{P}\right).$$

Conversely, when one wants to compute an arbitrary number (N) of discrete samples of one cycle of a continuous DTFT, $S_{1/T}(f)$, it can be done by computing the relatively simple DFT of $s_N[n]$, as defined above. In most cases, N is chosen equal to the length of non-zero portion of $s[n]$. Increasing N, known as *zero-padding* or *interpolation*, results in more closely spaced samples of one cycle of $S_{1/T}(f)$. Decreasing N, causes overlap (adding) in the time-domain (analogous to aliasing), which corresponds to decimation in the frequency domain. (see Sampling the DTFT) In most cases of practical interest, the $s[n]$ sequence represents a longer sequence that was truncated by the application of a finite-length window function or FIR filter array.

The DFT can be computed using a fast Fourier transform (FFT) algorithm, which makes it a practical and important transformation on computers.

See Discrete Fourier transform for much more information, including:

- transform properties
- applications
- tabulated transforms of specific functions

Summary

For periodic functions, both the Fourier transform and the DTFT comprise only a discrete set of frequency components (Fourier series), and the transforms diverge at those frequencies. One common practice (not discussed above) is to handle that divergence via Dirac delta and Dirac comb functions. But the same spectral information can be discerned from just one cycle of the periodic function, since all the other cycles are identical. Similarly, finite-duration functions can be represented as a Fourier series, with no actual loss of information except that the periodicity of the inverse transform is a mere artifact.

We also note that it is common in practice for the duration of $s(\bullet)$ to be limited to the period, P or N. But these formulas do not require that condition.

Duplex (telecommunications)

From Wikipedia, the free encyclopedia

(Redirected from Half-duplex)

This article includes a list of references, but **its sources remain unclear** because it has **insufficient inline citations**. Please help to improve this article by introducing more precise citations. *(September 2015)* *(Learn how and when to remove this template message)*

A **duplex** communication system is a point-to-point system composed of two connected parties or devices that can communicate with one another in both directions. Duplex systems are employed in many communications networks, either to allow for a communication "two-way street" between two connected parties or to provide a "reverse path" for the monitoring and remote adjustment of equipment in the field. There are two types of duplex communication systems: full-duplex (FDX) and half-duplex (HDX).

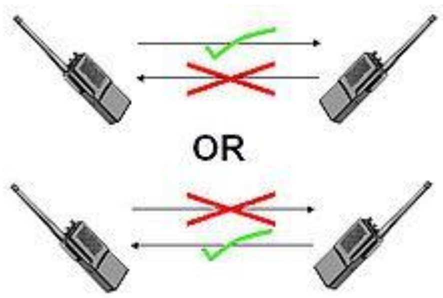
In a **full-duplex** system, both parties can communicate with each other simultaneously. An example of a full-duplex device is a telephone; the parties at both ends of a call can speak and be heard by the other party simultaneously. The earphone reproduces the speech of the remote party as the microphone transmits the speech of the local party, because there is a two-way communication channel between them, or more strictly speaking, because there are two communication channels between them.

In a **half-duplex** system, both parties can communicate with each other, but not simultaneously; the communication is one direction at a time. An example of a half-duplex device is a walkie-talkie two-way radio that has a "push-to-talk" button; when the local user wants to speak to the remote person they push this button, which turns on the transmitter but turns off the receiver, so they cannot hear the remote person. To listen to the other person they release the button, which turns on the receiver but turns off the transmitter.

Systems that do not need the duplex capability may instead use simplex communication, in which one device transmits and the others can only "listen". Examples are broadcast radio and television, garage door openers, baby monitors, wireless microphones, and surveillance cameras. In these devices the communication is only in one direction.

Contents

Half duplex



A simple illustration of a half-duplex communication system

A *half-duplex* (HDX) system provides communication in both directions, but only one direction at a time (not simultaneously). Typically, once a party begins receiving a signal, it must wait for the transmitter to stop transmitting, before replying.

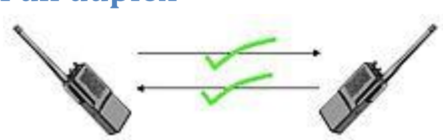
An example of a half-duplex system is a two-party system such as a walkie-talkie, wherein one must use "over" or another previously designated keyword to indicate the end of transmission, and ensure that only one party transmits at a time, because both parties transmit and receive on the same frequency. A good analogy for a half-duplex system would be a one-lane road with traffic controllers at each end, such as a two-lane bridge under re-construction. Traffic can flow in both directions, but only one direction at a time, regulated by the traffic controllers.

Half-duplex systems are usually used to conserve bandwidth, since only a single communication channel is needed, which is shared alternately between the two directions. For example, a walkie-talkie requires only a single frequency for bidirectional communication, while a cell phone, which is a full-duplex device, requires two frequencies to carry the two simultaneous voice channels, one in each direction.

In automatically run communications systems, such as two-way data-links, the time allocations for communications in a half-duplex system can be firmly controlled by the hardware. Thus, there is no waste of the channel for switching. For example, station A on one end of the data link could be allowed to transmit for exactly one second, then station B on the other end could be allowed to transmit for exactly one second, and then the cycle repeats.

In half-duplex systems, if more than one party transmits at the same time, a collision occurs, resulting in lost messages.

Full duplex



A simple illustration of a full-duplex communication system. Full-duplex is not common in handheld radios as shown here due to the cost and complexity of common duplexing methods, but is used in telephones, cellphones and cordless phones.

A *full-duplex* (FDX) system, or sometimes called *double-duplex*, allows communication in both directions, and, unlike half-duplex, allows this to happen simultaneously. Land-line telephone networks are full-duplex, since they allow both callers to speak and be heard at the same time, with the transition from four to two wires being achieved by a hybrid coil in a telephone hybrid. Modern cell phones are also full-duplex.[1]

A good analogy for a full-duplex system is a two-lane road with one lane for each direction. Moreover, in most full-duplex mode systems carrying computer data, transmitted data does not appear to be sent until it has been received and an acknowledgment is sent back by the other party. In this way, such systems implement reliable transmission methods.

Two-way radios can be designed as full-duplex systems, transmitting on one frequency and receiving on another; this is also called frequency-division duplex. Frequency-division duplex systems can extend their range by using sets of simple repeater stations because the communications transmitted on any single frequency always travel in the same direction.

Full-duplex Ethernet connections work by making simultaneous use of two physical twisted pairs inside the same jacket, which are directly connected to each networked device: one pair is for receiving packets, while the other pair is for sending packets. This effectively makes the cable itself a collision-free environment and doubles the maximum total transmission capacity supported by each Ethernet connection.

Full-duplex has also several benefits over the use of half-duplex. First, there are no collisions so time is not wasted by having to retransmit frames. Second, full transmission capacity is available in both directions because the send and receive functions are separate. Third, since there is only one transmitter on each twisted pair, stations (nodes) do not need to wait for others to complete their transmissions.

Some computer-based systems of the 1960s and 1970s required full-duplex facilities, even for half-duplex operation, since their poll-and-response schemes could not tolerate the slight delays in reversing the direction of transmission in a half-duplex line.

Channel access method

In telecommunications and computer networks, a **channel access method** or **multiple access method** allows several terminals connected to the same multi-point transmission medium to transmit over it and to share its capacity.[1] Examples of shared physical media are wireless networks, bus networks, ring networks and point-to-point links operating in half-duplex mode.

A channel access method is based on multiplexing, that allows several data streams or signals to share the same communication channel or transmission medium. In this context, multiplexing is provided by the physical layer.

A channel access method is also based on a multiple access protocol and control mechanism, also known as media access control (MAC). Media access control deals with issues such as addressing, assigning multiplex channels to different users, and avoiding collisions. Media access control is a sub-layer in the data link layer of the OSI model and a component of the link layer of the TCP/IP model.

Contents

Fundamental types of channel access schemes

These numerous channel access schemes which generally fall into the following categories:[2][3][1]

Frequency-division multiple access (FDMA)

The frequency-division multiple access (FDMA) channel-access scheme is based on the frequency-division multiplexing (FDM) scheme, which provides different frequency bands to different data-streams. In the FDMA case, the data streams are allocated to different nodes or devices. An example of FDMA systems were the first-generation (1G) cell-phone systems, where each phone call was assigned to a specific uplink frequency channel, and another downlink frequency channel. Each message signal (each phone call) is modulated on a specific carrier frequency.

A related technique is wavelength division multiple access (WDMA), based on wavelength-division multiplexing (WDM), where different datastreams get different colors in fiber-optical communications. In the WDMA case, different network nodes in a bus or hub network get a different color.

An advanced form of FDMA is the orthogonal frequency-division multiple access (OFDMA) scheme, for example used in 4G cellular communication systems. In OFDMA, each node may use several sub-carriers, making it possible to provide different quality of service (different data rates) to different users. The assignment of sub-carriers to users may be changed dynamically, based on the current radio channel conditions and traffic load.

Time division multiple access (TDMA)

The time division multiple access (TDMA) channel access scheme is based on the time-division multiplexing (TDM) scheme, which provides different time-slots to different data-streams (in the TDMA case to different transmitters) in a cyclically repetitive frame structure. For example, node 1 may use time slot 1, node 2 time slot 2, etc. until the last transmitter. Then it starts all over again, in a repetitive pattern, until a connection is ended and that slot becomes free or assigned to another node. An advanced form is Dynamic TDMA (DTDMA), where a scheduling may give different time sometimes but some times node 1 may use time slot 1 in first frame and use another time slot in next frame.

As an example, 2G cellular systems are based on a combination of TDMA and FDMA. Each frequency channel is divided into eight timeslots, of which seven are used for seven phone calls, and one for signalling data.

Statistical time division multiplexing multiple-access is typically also based on time-domain multiplexing, but not in a cyclically repetitive frame structure. Due to its random character it can be categorised as statistical multiplexing methods, making it possible to provide dynamic bandwidth allocation. This requires a media access control (MAC) protocol, i.e. a principle for the nodes to take turns on the channel and to avoid collisions. Common examples are CSMA/CD, used in Ethernet bus networks and hub networks, and CSMA/CA, used in wireless networks such as IEEE 802.11.

Code division multiple access (CDMA)/Spread spectrum multiple access (SSMA)

The code division multiple access (CDMA) scheme is based on spread spectrum, meaning that a wider radio spectrum in Hertz is used than the data rate of each of the transferred bit streams, and several message signals are transferred simultaneously over the same carrier frequency, utilizing different spreading codes. The wide bandwidth makes it possible to send with a very poor signal-to-noise ratio of much less than 1 (less than 0 dB) according to the Shannon-Heartly formula, meaning that the transmission power can be reduced to a level below the level of the noise and co-channel interference (cross talk) from other message signals sharing the same frequency.

One form is direct sequence spread spectrum (DS-CDMA), used for example in 3G cell phone systems. Each information bit (or each symbol) is represented by a long code sequence of several pulses, called chips. The sequence is the spreading code, and each message signal (for example each phone call) uses a different spreading code.

Another form is frequency-hopping (FH-CDMA), where the channel frequency is changing very rapidly according to a sequence that constitutes the spreading code. As an example, the Bluetooth communication system is based on a combination of frequency-hopping and either CSMA/CA statistical time division multiplexing communication (for data communication applications) or TDMA (for audio transmission). All nodes belonging to the same user (to the same virtual private area network or piconet) use the same frequency hopping sequence synchronously, meaning that they send on the same frequency channel, but CDMA/CA or TDMA is used to avoid collisions within the VPAN. Frequency-hopping is used to reduce the cross-talk and collision probability between nodes in different VPANs.

Subdivisions of FH-CDMA are "fast hopping" where the frequency of hopping is much higher than the message frequency content and "slow hopping" where the hopping frequency is comparable to message frequency content. The subdivision is necessary as they are considerably different.

Space division multiple access (SDMA)

Space-division multiple access (SDMA) transmits different information in different physical areas. Examples include simple cellular radio systems and more advanced cellular systems which use directional antennas and power modulation to refine spatial transmission patterns.

Power division multiple access (PDMA)

Power-division multiple access (PDMA) scheme is based on using variable transmission power between users in order to share the available power on the channel. Examples include multiple SCPC modems on a satellite transponder, where users get on demand a larger share of the power budget to transmit at higher data rates.[4]

List of channel access methods

Circuit mode and channelization methods

The following are common circuit mode and channelization channel access methods:

- ***Frequency-division multiple access (FDMA)***, based on frequency-division multiplexing (FDM)
  - Wavelength division multiple access (WDMA)
  - Orthogonal frequency-division multiple access (OFDMA), based on Orthogonal frequency-division multiplexing (OFDM)
  - Single-carrier FDMA (SC-FDMA), a.k.a. linearly-precoded OFDMA (LP-OFDMA), based on single-carrier frequency-domain-equalization (SC-FDE).
- ***Time-division multiple access (TDMA)***, based on time-division multiplexing (TDM)
  - Multi-Frequency Time Division Multiple Access (MF-TDMA)
- ***Code division multiple access (CDMA)***, a.k.a. Spread spectrum multiple access (SSMA)
  - Direct-sequence CDMA (DS-CDMA), based on Direct-sequence spread spectrum (DSSS)

    - Frequency-hopping CDMA (FH-CDMA), based on Frequency-hopping spread spectrum (FHSS)
    - Orthogonal frequency-hopping multiple access (OFHMA)
    - Multi-carrier code division multiple access (MC-CDMA)
- ***Space-division multiple access (SDMA)***
- Power-division multiple access (PDMA)

Packet mode methods

The following are examples of packet mode channel access methods:[1]

- ***Contention based random multiple access methods***
    - Aloha
    - Slotted Aloha
    - Multiple Access with Collision Avoidance (MACA)
    - Multiple Access with Collision Avoidance for Wireless (MACAW)
    - Carrier sense multiple access (CSMA)
    - Carrier sense multiple access with collision detection (CSMA/CD) - suitable for wired networks
    - Carrier sense multiple access with collision avoidance (CSMA/CA) - suitable for wireless networks
        - Distributed Coordination Function (DCF)
    - Carrier sense multiple access with collision avoidance and Resolution using Priorities (CSMA/CARP)
    - Carrier Sense Multiple Access/Bitwise Arbitration (CSMA/BA) Based on constructive interference (CAN-bus)
- ***Token passing***:
    - Token ring
    - Token bus
- ***Polling***
- ***Resource reservation (scheduled) packet-mode protocols***
    - Dynamic Time Division Multiple Access (Dynamic TDMA)
    - Packet reservation multiple access (PRMA)
    - Reservation ALOHA (R-ALOHA)

Duplexing methods

Where these methods are used for dividing forward and reverse communication channels, they are known as duplexing methods, such as:

- Time division duplex (TDD)
- Frequency division duplex (FDD)

Hybrid channel access scheme application examples

Note that hybrids of these techniques can be - and frequently are - used. Some examples:

- The GSM cellular system combines the use of frequency division duplex (FDD) to prevent interference between outward and return signals, with FDMA and TDMA to allow multiple handsets to work in a single cell.

- GSM with the GPRS packet switched service combines FDD and FDMA with slotted Aloha for reservation inquiries, and a Dynamic TDMA scheme for transferring the actual data.
- Bluetooth packet mode communication combines frequency hopping (for shared channel access among several private area networks in the same room) with CSMA/CA (for shared channel access inside a medium).
- IEEE 802.11b wireless local area networks (WLANs) are based on FDMA and DS-CDMA for avoiding interference among adjacent WLAN cells or access points. This is combined with CSMA/CA for multiple access within the cell.
- HIPERLAN/2 wireless networks combine FDMA with dynamic TDMA, meaning that resource reservation is achieved by packet scheduling.
- G.hn, an ITU-T standard for high-speed networking over home wiring (power lines, phone lines and coaxial cables) employs a combination of TDMA, Token passing and CSMA/CARP to allow multiple devices to share the medium.

Definition within certain application areas

Local and metropolitan area networks

In local area networks (LANs) and metropolitan area networks (MANs), multiple access methods enable bus networks, ring networks, hubbed networks, wireless networks and half duplex point-to-point communication, but are not required in full duplex point-to-point serial lines between network switches and routers, or in switched networks (logical star topology). The most common multiple access method is CSMA/CD, which is used in Ethernet. Although today's Ethernet installations typically are switched, CSMA/CD is utilized anyway to achieve compatibility with hubs.

Satellite communications

In satellite communications, multiple access is the capability of a communications satellite to function as a portion of a communications link between more than one pair of satellite terminals concurrently. Three types of multiple access presently used with communications satellites are code-division, frequency-division, and time-division multiple access.

Switching centers

In telecommunication switching centers, multiple access is the connection of a user to two or more switching centers by separate access lines using a single message routing indicator or telephone number.

Classifications in the literature

Several ways of categorizing multiple-access schemes and protocols have been used in the literature. For example, Daniel Minoli (2009)[5] identifies five principal types of multiple-access schemes: FDMA, TDMA, CDMA, SDMA, and Random access. R. Rom and M. Sidi (1990)[6] categorize the protocols into *Conflict-free access protocols*, *Aloha protocols*, and *Carrier Sensing protocols*.

The Telecommunications Handbook (Terplan and Morreale, 2000)[7] identifies the following MAC categories:

- Fixed assigned: TDMA, FDMA+WDMA, CDMA, SDMA
- Demand assigned (DA)
    - Reservation: DA/TDMA, DA/FDMA+DA/WDMA, DA/CDMA, DA/SDMA
    - Polling: Generalized polling, Distributed polling, Token Passing, Implicit polling, Slotted access
- Random access (RA): Pure RA (ALOHA, GRA), Adaptive RA (TRA), CSMA, CSMA/CD, CSMA/CA