NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Electrical & Electronics Engineering

Public Switching Telephone Network

GRADUATION PROJECT
EE – 400

Student: Naveed Akhtar (980689)

Supervisor: Prof. Dr. Fakhraddin Mamedov

Nicosia - 2002

Dedicated to my beloved mother whose love and kindness has made me "who I am".

ACKNOWLEDGEMENTS

I am thankful to my supervisor Prof Dr. Fakhreddm Mamedov for his supervision and constructive criticism in the development of this project. He has made this achievement possible for me. He guided me too much and I can say without Him this project would not have been possible. His words of encouragement kept us going, and under his supervision it is a writing good experience for me. I took 3 subjects from him about communication and get good grads because under hid guidance, I successfully overcome many difficulties and learn a lot about Communication. His way of teaching is so kind that I could under stand whole lecture at the moment. I chose this project because he has good knowledge on communication.

Special thank to my friend Imran Ahmed for his kind help he has a great importance in my heart, and also thankful to my friends Umer and Osama, who helped me in this project. I also want thank to my friends: Asif, Sohail, Tahir, Nabeel, Yousaf and Rihan, being with them make my 4 years in NEU full of fun and enjoy.

At the end I want to thank to my family and especially paying regards to my parents, who help me on every stage of my life where ever I need, and they pick me up about studies and in all matters of life, and also special regard to my father who spent 25 years in abroad only for his children to give them better future. It is only because of my parent's prayers and endless efforts that I am completing my degree. I whish they always be happy.

ABSTRACT

In daily life PSTN has a great importance and used in every important field like in official work, Army field and field of science.

A telecommunication system can take many different forms PSTN has a very important role in our life.

Chapter 1 includes the history and background of telecommunication and also have information about major elements and types of communication device, systems and importance of PSTN in communication.

Chapter 2 presents five major principal: sampling, quantizing, electrical representation of PCM signals, coding and companding , demodulation for PCM systems. This theoretical material is used it be representation first order PCM system. The functional and timing diagram systems are presented.

Chapter 3 briefly describe telephone that how telephone works and how the systems layout is. It includes the three basic and major systems: signaling, transmission and switching.

Chapter 4 briefly described the failure of PSTN and which factors would be involved to fail the Public Switching Telephone Network (PSTN).

TABLE OF CONTANTS

AKNOWLEDGEMENT i

ABSTRACTS ii

1. INTRODUCTION TO TELECUMMUNICATION & PSTN

1.2. History of Telecommunication 1

1.2.1. Networks 3

1.3. Public Switching Telephone Network (PSTN) 4

1.4. Telecommunication Concept 6

1.4.1. Analogue and Digital Networks 7

1.4.2. Circuit-Switching and Packet-Switching 7

1.4.3. Performance 8

2. THE FUNCTION BLOCK OF THE PCM SYSTEM 9

2.1. Sampling 11

2.2. Quantizing 13

2.3. Coding & Companding 14

2.4. Electrical Representation of PCM signals 20

2.5. Demodulation 22

3. TELEPHONE NETWORK 24

3.1. Signal Functions 25

3.1.1. Subscriber Loop Signaling 25

3.1.2. Interoffice Signaling 26

3.1.3. Common Channel Interoffice Signaling 27

3.1.4. Telephone Numbering Plan 29

3.1.5. Local Loop Signaling Design 29

3.2. Transmission 30

3.3. Open Wire and Paired Cable 31

3.3.1. Loading Coils 32

3.3.2. Multiplexing 33

3.3.3. Analog Multiplexing 34

3.3.4. Digital Multiplexing 35

3.4. Coaxial Cable 35

3.5. Micro Wave Terrestrial System 36
3.6. Communication Satellites 37
3.6.1. Frequency Band 39
3.6.2. Delay 41
3.6.3. Access 42
3.6.4. Capacity 43
3.7. Optical Fiber 43
3.8. Transmission Capacities 46
3.9. Echo Elimination 47
3.10. Switching 48
3.10.1. Services Evolution 48
3.10.2. Technology evolution 50
3.11. Network Switching 52
3.12. Approaches to Switching 54
3.12.1. Space and Time Division 54
3.12.2. Technologies 55
3.13. Space Division Switching 55
3.14 Control Method 56
3.14.1 Direct progressive Control 56
3.14.2 Register Progressive Control 59
3.14.3 Common control 59
3.14.4 Stored Program Control 60
3.15 Step-by-Step Switching Control 61
3.15.1 The Strowger Switch 62
3.15.2 Switching System 62
4. SOURCE OF FALIURE IN THE PSTN 65
4.1. Failure Classifications 66
4.2. Analyses Procedure 67
4.3. Findings 70
4.4. Observations 71
4.5. Why So Reliable? 72
4.5.1. Reliable software 73
4.5.2. Dynamic rerouting 73

4.6. Loose Coupling 74
4.7 Human Intervention 75
CONCLUSION 76
REFERENCE 77

Chapter 1

Introduction to Telecommunication and PSTN

1 Telecommunication

Telecommunications Communications over a distance using technology to overcome that distance. It usually means the transmission of words, sounds, pictures, or data in the form of electronic signals or impulses, sent either as an individual message between two parties or as a broadcast to be received at many locations. While broadcasting is far removed from private communications, a new range of one-to-one communication.

Services (including video-on-demand and other personal information and entertainment services provided over cable networks and so-called "webcasting" over the Internet) will blur the current clear distinction between the two.

Since its invention by Alexander Graham Bell in 1876, the telephone has become the most familiar form of telecommunications. More recently, a range of computer-based telecommunication services has supplemented voice telephony. These have become popular through the Internet and World Wide Web-vast computer networks that provide many people with the means to exchange information.

1.2 History of Telecommunication

It is now taken for granted in developed nations that by pressing a few buttons people can talk to family, friends, or business associates across the world. The technology that has led to one of the most complex creations of the 20th century—the telephone network—has evolved over the past hundred years or so.

The first electrical means of communication was not the telephone, however, but the telegraph, which allowed messages, sent in code (usually Morse Code) to be received and printed at a distant location. The age of commercial telegraphy dawned in 1839 when the British pioneers William Fothergil Cooke and Charles Wheatstone opened their line alongside the main railway route running west from London. Samuel Morse devised a technically simpler system of telegraphy in 1843, and after this the spread of telegraph networks was rapid, with routes spreading across most of the countries of the Old and New Worlds and then beneath the oceans that separated them. By 1930 nearly 650,000 km (400,000 mi) of undersea cables had been laid, linking the economic, political, military, and cultural institutions of the world.

An even greater breakthrough was made in 1876, when Alexander Graham Bell made the first telephone call to his assistant with the words "Mr Watson, come here, I want you". Bell's invention sparked a series of innovations, ultimately culminating in today's information superhighway. Key steps along the way were:

In 1889 Almon Strowger developed an automatic switching system that could set up a telephone call without intervention by a human operator. Strowger's motivation for this invention was to prevent his calls being diverted to a business competitor by his local operator. The impact of the invention was much wider as it provided the basis for the current telephone network.

In 1901 Guglielmo Marconi demonstrated that radio waves could be used to transmit information over long distances when he sent a radio message across the Atlantic. Radio is still one of the key transmission media today, and is the basis of many mobile services.

In 1947 William Shockley, John Bardeen, and Walter Brattain invented the transistor. This enabled the electronics revolution to take place and provided the basis for a computerized, rather than mechanical, telecommunications network.

In 1965 Charles Kao put forward the theory that information could be carried using optical fibers. These have subsequently been developed to provide a means of carrying huge amounts of information at very high speed. Optical fibers form the backbone of the global transmission network.

The modem telephone network can be viewed as a globally distributed machine that operates as a single resource. Much of it uses interconnected computers. The network that most people use to carry voice traffic can also be used to transfer data in the form of pictures, text, and video images.

1.2.1 Networks

Despite being very complex, global telecommunications service is comprised of a few basic network components, which are: (1) user equipment—telephones, computers, and all the other devices that provide a means of accessing the network; (2) the access network—users are connected to the main network by wire line or radio links; (3) the main network—copper wire, microwave radio, and optical fiber cables connecting all the nodes of the global network; (4) transmission equipment—the means by which huge volumes of information (there are many millions of telephone and data calls made every second) are carried over the network; and *(5)* switching equipment—the hierarchy of local, long-distance, and international switches that allow any user of the network to connect to any other user. Each of these components has to consist of a combination of hardware and software.

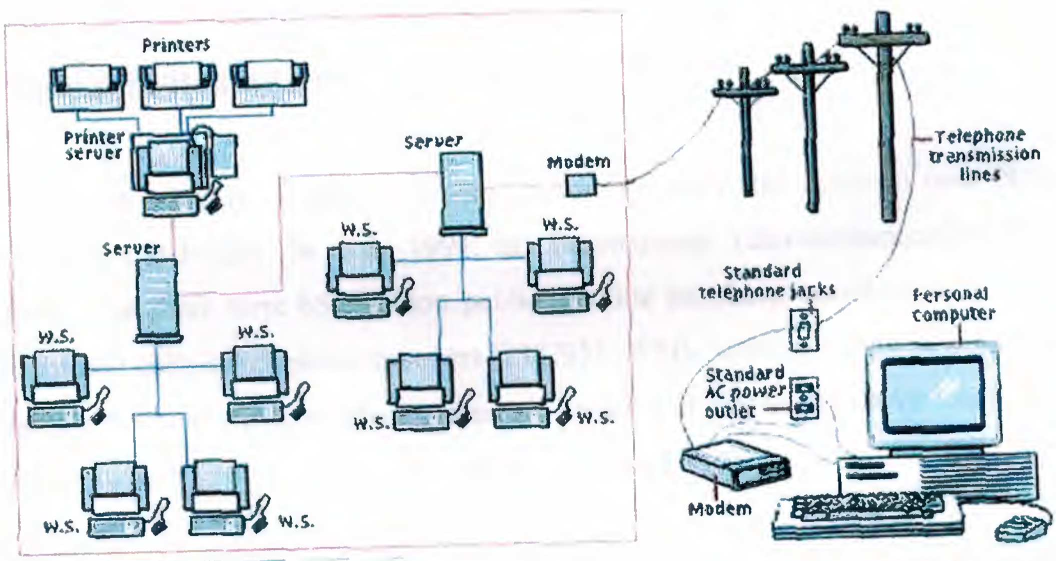


Figure 1.1

a) Hardware

This usually covers items such as telephones, transmitters, cables, interface devices, switches, and computers. In the past, telecommunications have relied heavily on hardware, such as dedicated switching elements, and on the logic providing its control functions. A situation is now developing in which more of the system relies on elements operating under computer (software) control. Because this software can be upgraded, this makes it easy to add new, enhanced functionality later.

b) Software

This is code that instructs a computer or network device. Until the I980s, most of the operational instructions used by a telecommunications network were hard-wired or pre-set. The advent of digital systems and data networks has led to a much wider range of network services. Software solutions are well suited to the complexity and flexibility inherent in these services.

1.3 Public Switching Telephone Network (PSTN):

The PSTN is a highly integrated communications network that connects over 70% of the world's inhabitants. In early 1994, the International Telecommunications Union estimated that there were 650 million public landline telephone numbers, as compared to 30 million cellular telephone numbers [1TU931. While landline telephones are, being added at a 3% rate, wireless subscriptions are growing at greater than a 50% rate. Every telephone in the world is given calling access over the PSTN.

Each country is responsible for the regulation of the PSTN within its borders. Over time, some government telephone systems have become privatized by corporations which provide local and long distance service for profit.

In the PSTN, each city or a geographic grouping of towns is called a local access and transport area (LATA). Surrounding LATAs are connected by a company called a local exchange carrier (LEC). A LEC is a company that provides intralata telephone service.

and may be a local telephone company, or may be a telephone company that is regional in scope. A long distance telephone company collects toll fees to provide connections between different LATAs over its long distance network. These companies are referred to as interexchange carriers (IXC), and own and operate large fiber optic and microwave radio networks which are connected to LECs throughout a country or continent. (Figure 1.1) is a simplified illustration of a local telephone network, called a local exchange. Each local exchange consists of a central office (CO) which provides Figure (1.2) is a simplified illustration of a local telephone network, called local exchange. Each local exchange consists of a central office (CO) which provides PSTN connection to the customer premises equipment (CPE) which may he an individual phone at a residence or a private branch exchange (PBX) at a place of business. The CO may handle as many as a million telephone connections.

The CO is connected to a tandem switch 'which in turn connects the local exchange to the PSTN. The tandem switch physically connects the local telephone network to the point of presence (POP) of trunked long distance lines provided by one or more IXCs [17Pec92]. Sometimes IXCs connect directly to the CO switch to avoid local transport charges levied by the LEC.

Figure (1.2) also shows how a PBX may he used to provide telephone connections throughout a building or campus. A PBX allows an organization or entity to provide internal calling and other in—building services (which do not involve the LEC). as well as private networking between other organizational sites (through leased lines from EEC and IXC providers), in addition to conventional local and long distance services which pass through the CO. Telephone connections within a PBX are maintained by the private owner, whereas connection of the PBX to the CO is provided and maintained by the LEC.

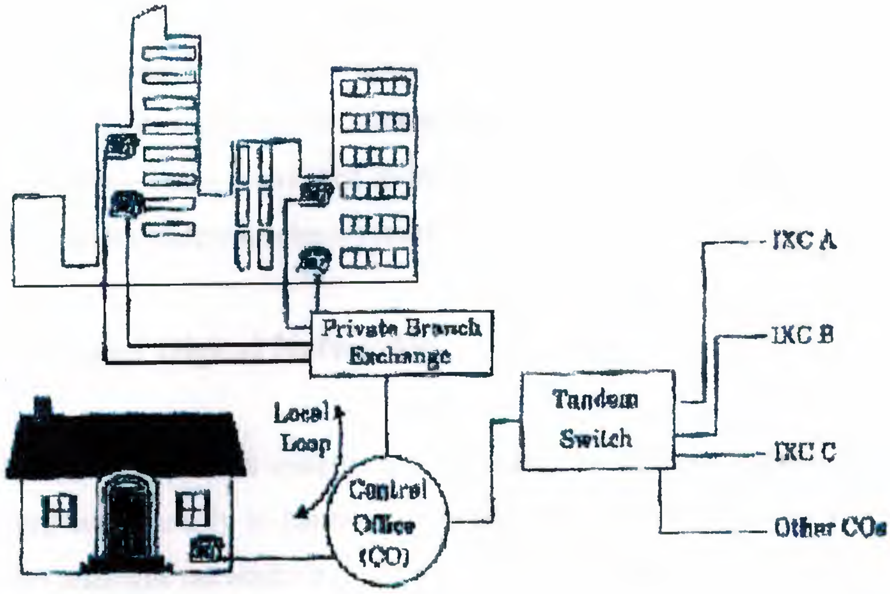


Figure 1.2

Since the invention of the telephone, the public switch telephone network (PSTN) has grown proportionately with the increase demand to communicate. Switching services beyond metropolitan areas were soon developed increasing the size and complexity of the central office. New methods of switching were required to interconnect central offices through the use of interoffice trunks and tandem trunks as shown in figure. When the call is made outside the local area, they are routed through Toll Trunk and Toll Center.

1.4 Telecommunication Concept:

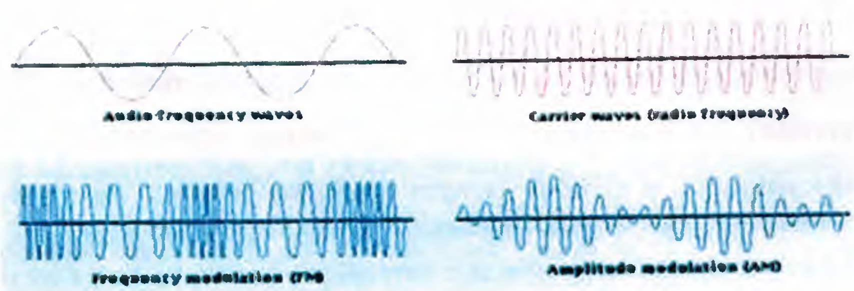


Figure 1.3

There are several ways of carrying information between senders and users. The options chosen should reflect the type of communication required. For instance, humans compensate for noise and transmission errors when they talk to each other. Unexpected delays or echoes cause problems in understanding, however. Computers have the reverse characteristics—being tolerant of short delays and less so of transmission errors. The following concepts underpin telecommunications networks.

1.4.1 Analogue and Digital Networks:

Many older telecommunications systems are analogue; the electrical signals conveying information vary continuously in harmony with the sounds they represent. The quality of speech across analogue networks is determined by the amount of the speech spectrum that could be carried. Around 3 kHz was accepted as a reasonable compromise of cost and quality for normal telephone calls.

The alternative way of transmitting information is with a straightforward electrical signal that is either on or off, as with Morse's telegraph. Computers also communicate with discrete, digital (on/oft) signals, and while these can be converted to tones for transmission over analogue communications, it makes more sense to send them back in their original digital form. Speech and other analogue communications can readily be converted into digital form, and back to analogue *(see* Digital-to-Analogue Converter and Analogue-to-Digital Converter). Most telecommunications networks today are "integrated" digital systems, ideally suited to computer networking and other multimedia applications such as speech (voice), data, text, fax, and video.

1.4.2 Circuit-Switching and Packet-Switching

The distinguishing feature of circuit-switching is that an end-to-end connection is set up between the communicating parties, and is maintained until the communication is complete. The public switched telephone network (PSTN) is a familiar example of a circuit-switched network.

Communication between computers, or between computers and terminals, always involves the transfer of data in blocks rather than continuous data streams. Packet-switching exploits the fact that data blocks can be transferred between terminals without setting up an end-to-end connection through the network. Instead they are transmitted on a link-by-link basis, being stored temporarily at each switch en route where they queue for transmission on an appropriate outgoing link. Routing decisions are based on addressing information contained in a "header" appended to the front of each data block. The term "packet" refers to the header plus data block.

Congestion and Blocking In a packet-switched network, packets compete dynamically for the network's resources (buffer storage, processing power, transmission capacity). A switch accepts a packet from a terminal largely in ignorance of what resources the network will have available to handle it. There is always the possibility, therefore, that a network will admit more traffic than it can actually carry with a corresponding degradation in service. Controls are therefore needed to ensure that such congestion does not arise too often and that the network recovers gracefully when it does.

In a circuit-switched network, the competition for resources takes the form of "blocking". This means that one user's call may prevent another user from getting access. Since the user reserves the circuit—irrespective of what is sent—for the duration of the user's call, no one else has any form of access until the call is cleared. Traditional circuit-switched networks are designed to balance the amount of equipment deployed against a reasonable level of access for the users of that network.

1.4.3 Performance

A circuit-switched network, such as the PSTN, provides end-to-end connections on demand, as long as the necessary network resources are available. The connection's end-to-end delay is usually small and always constant and other users cannot interfere with the quality of communication. In contrast, in a packet-switched network, packets queue for transmission at each switch. The cross-network delay is therefore variable as it depends on the volume of traffic encountered en route and if it exceeds a certain level, system performance can be badly impaired.

Chapter 2

THE FUNCTION OF THE PCM SYSTEM

2 Introduction to PCM

The statement above gives some idea about the basic processes in pulse code modulation. Here we shall give these processes their right names.

The process of choosing measuring points on the analogue speech curve is called sampling. The measurement values are called samples. When sampling. We take the first step towards a digital representation of the speech signal as the chosen sampling instants give us the time coordinates of the measuring points.

The amplitudes of the samples can assume each value in the amplitude range of the speech signal. When measuring the sample amplitudes we have to round off for practical reasons. In the rounding-off process, or the quantizing process, all sample amplitudes between two marks on the scale will be given the same quantized value. The number of quantized samples is discrete as we have only a discrete number of marks on our scale.

Each quantized sample is then represented by the number of the scale mark, i.e. we know now the coordinates on the amplitude axis of the samples.
The processes of sampling and quantizing yield a digital representation of the original speech signal, but not in a form best suited to transmission over a line or radio path. Translation to a different form of signal is required. This process is known as encoding. Most often the sample values are encoded to binary form, so that each sample value is represented by a group of binary elements. Typically, a quantized sample can assume one of 256 values. In binary form the sample will be represented by a group of 8 elements. This group is in the following called a PCM word. For transmission purposes the binary values 0 and I can be taken as corresponding to the absence and presence of an electrical pulse.

On the transmission line the pulses in the PCM words will become gradually more distorted. However, as long as it is possible to distinguish between the absence and the presence of a pulse, no information loss has occurred. If the pulse train is regenerated, i.e. badly distorted pulses are replaced by fresh pulses at suitable Intervals, the information can be transmitted long distances with practically no distortion at all. This is one of the advantages of digital transmission over analogue transmission; the information is contained in the existence or not of a pulse rather than in the form of the pulse.

In our picture of the graph and the table this is analogous to the fact that the information in the table is not affected if the digits are badly written as long as they are legible. But if the graph is badly drawn, loss of information is inevitable.

On the receiving side the PCM words are decoded. i.e. they are translated back to quantized samples. The analogue speech signal is then reconstructed by interpolation between the quantized samples. There is a small difference between the analogue speech signal on the receiving side and the corresponding signal on the transmitting side due to the rounding off of the speech samples. This difference is known as quantizing distortion.

The function blocks in the pulse code modulation process are shown in figure 2.11.

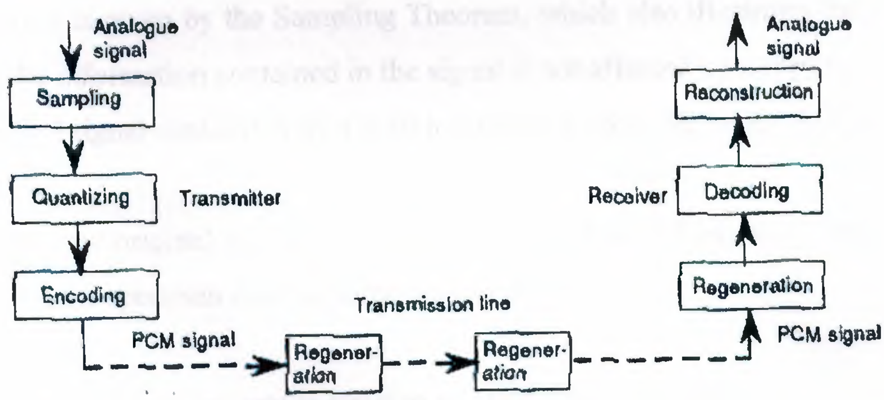


Figure 2.1 Pulse code modulation Function Block

2.1 Sampling

In the practical electrical meaning, to sample is to take instaneous values of the analogue signal at equal time intervals. See figure 2.2.

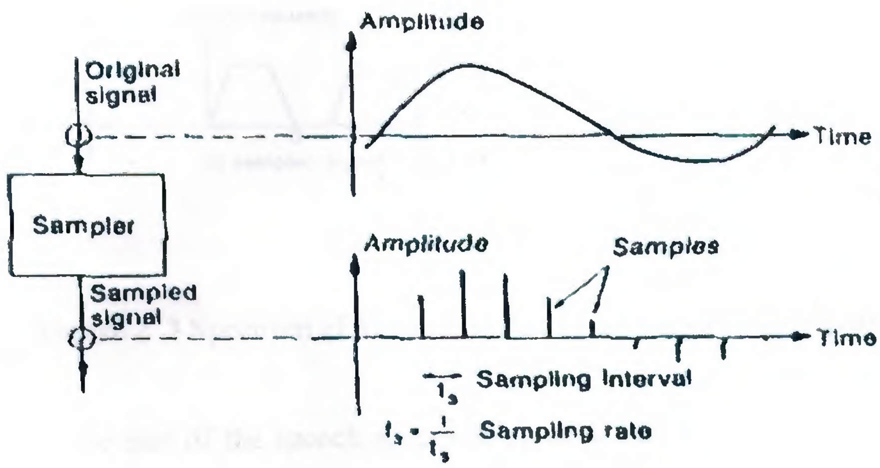


Figure 2.2. The sampling process

The sampled signal is a train of pulses, whose envelope is the original signal. Now, what should be the sampling rate, i.e. the number of samples per second? The answer to this question is given by the Sampling Theorem, which also illustrates the fundamental fact that the information contained in the signal is not affected by sampling:

The sampled signal contains within it all information about the original signal if:

- the original signal is band limited, i.e. it has no frequency components in its spectrum beyond some frequency B

- the sampling rate is equal to or greater than twice B, i.e. $f_s \geq 2B$.

The sampling theorem is illustrated in figure 2-3. Obviously, the spectrum of the sampled signal contains the spectrum of the original signal, i.e. no information loss has occurred.

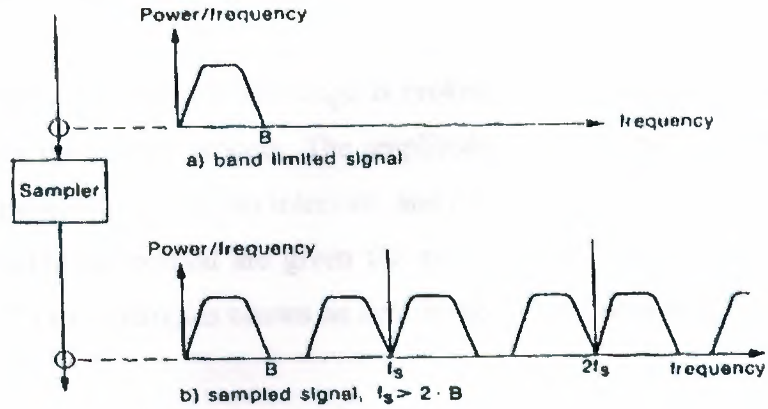


Figure 2 .3 Spectrum of a) band-limited signal b) sampled signal

In telephony, the part of the speech spectrum between 300 and 3400 Hz is used. The human speech spectrum extends from a lowest frequency of some 100 Hz up to very high audio frequencies. The telephone set reduces this frequency range, but not enough at high frequencies so in order to come below this band limit at 3400 Hz, the speech signal must be low-pass -filtered before sampling.

A sampling rate of 8000 Hz is used for PCM systems in telephony. This rate is somewhat higher than twice the highest frequency in the band, 3400 Hz, due to difficulties in making low-pass filters steep enough.

The sampled signal is often said to be pulse amplitude modulated as it consists of a train of pulses, whose amplitudes have been modulated by the original signal. Pulse Amplitude Modulation (PAM) is an analogue pulse modulation method as the amplitudes of the pulses may vary continuously in accordance with the original signal variations.

The relative simplicity of PAM systems makes them attractive for some telephony applications. However, PAM is unsuitable for transmission over long distances owing to the difficulty of pulse regeneration with sufficient accuracy, which is important as the PAM pulses contain the information, in the pulse form.

2.2. Quantizing

The continuous pulse amplitude range is broken down to a finite number of amplitude values in the quantizing process. The amplitude values in the quantizing process. The amplitude range is divided into intervals, and all samples whose amplitudes fall into one specific quantizing interval are given the same output amplitude. See figure 1-4. The rounding off of the samples causes an irretrievable error, squinting distortion, in the signal.

This voluntary sacrifice, which can be brought down to suitable low limits by making the number of permitted amplitude levels large enough, is accepted because it makes error-free transmission possible by only having a discrete number of amplitudes.
In figure 2-4, the quantizing distortion is independent of sample amplitude. This means that a loud talker and quiet talker let a listener hear the same quantizing distortion. Relative to the speech levels, the quiet talker generates much more distortion than the loud talker. Furthermore, a statistical analysis shows that for an Individual talker, small amplitudes are much more probable than large ones.

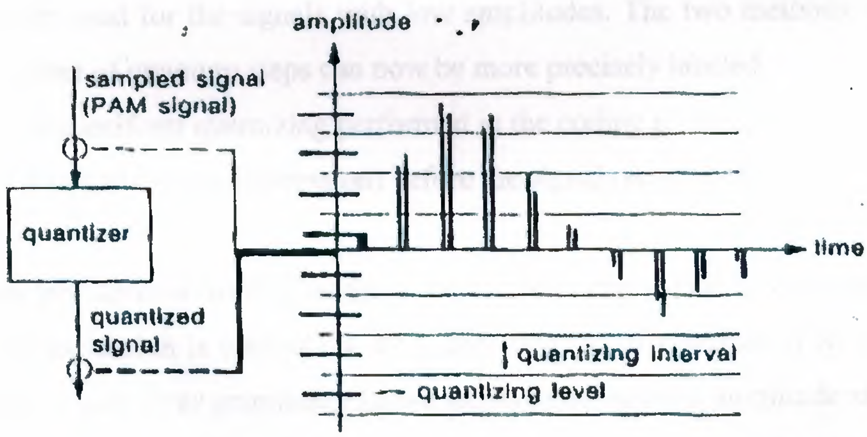


Figure2.4. The quantizing process

2.3. Coding & Companding

Practical PC systems use seven-and eight-level binary codes, or

$$2^7 = 128 \text{ quantum steps}$$

$$2^8 = 256 \text{ quantum steps}$$

Two methods are used to reduce the quantum steps to 128 or 256 without sacrificing fidelity. These are nonuniform quantizing steps and companding before quantizing, followed by uniform quantizing. Unlike data transmission, in speech transmission there is a much greater likelihood of encountering signals of small amplitudes than those of larger amplitudes.

A secondary but equally important aspect is that coded signals are designed to convey maximum Information, considering that all quantum steps (meanings or characters) will have an equally probable occurrence (i.e., the signal-level amplitude is assumed to follow a uniform probably distribution between 0 and ± the maximum voltage of the channel). To circumvent the problem of nonequiprobability of signal level for voice signals, specifically, that lower - level signal are more probable than higher-level signals, larger quantum steps are used for the larger-amplitude portion of the signal, and finer steps are used for the signals with low amplitudes. The two methods of reducing the total number of quantum steps can now be more precisely labeled:

- Nonuniform quantizing performed in the coding process.
- Companding (compression) before the signals enter the coder,

Which now performs uniform quantizing on the resulting signal before coding. At the receive end, expansion is carried out after decoding. Most practical PCM systems use companding to give finer granularity (more steps) to the smaller amplitude signals. This is instantaneous companding, as compared to the syllabic companding used in analog carder telephony. Compression Imparts more gain to lower amplitude signals. The compression and later expansion functions are logarithmic and follow one of two laws, the A law or the "mu" (μ) law.

A common expression used in dealing with the "quality' of a PCM signal is the signal-to-distortion ratio (expressed in decibels). Parameters A and μ. determine the range over which the signal-to-distortion ratio Is comparatively constant. This is the dynamic range. Using a p. of 100 can provide a dynamic range of 40 dB of relative linearity in the signal-to-distortion ratio.

In actual PCM systems, the companding circuitry does not provide an exact replica of the logarithmic curves shown. The circuitry produces approximate equivalents using a segmented curve, and each segment is linear. The more segments the curve has, the more it approaches the true logarithmic curve desired. Such a segmented curve is shown in Figure 1-5. If μ law were implemented using a seven (height)-segment linear approximate equivalent, it would appear as shown in Figure 1-5. Thus on coding, the first three coded digits would indicate the segment number (e.g. $2^3 = 8$). Of the seven-digit code, the remaining four digits would divide each segment into 16 equal parts to identify further the exact quantum step (e.g., $2^4 = 16$) For small signals, the companding improvement.

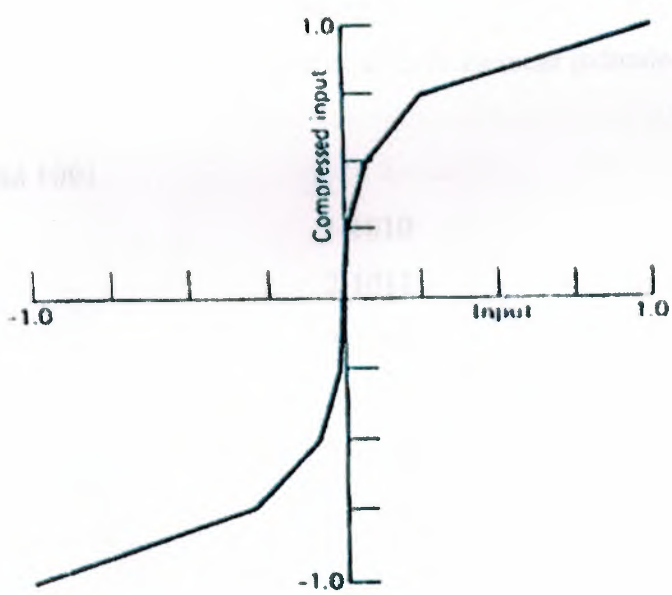


Figure 2.5. Seven-segment linear approximate of the logarithmic curve for μ law (μ = 100)is approximately, A law: 24 dB μ law: 30 dB

using a seven-level code. These values derive from the equation of companding improvement or coding in PCM systems utilizes straightforward binary codes. Examples of such coding are shown in Figure I-5a, which is expanded in Figure 9.7, and in Figure 9.8, which is expanded in Figure 1 .6.b showing a number of example code levels.

The coding process is closely related to quantizing. In practical systems, whether the A law or the μ law is used, quantizing employees segmented equivalents of the companding curve (Figures 1-6 andl-8), as discussed earlier. Such segmenting is a handy aid to coding. Consider the European 30 + 2 PCM system, which uses a 13-segment approximation of the A-law curve (Figure 1-6). The first code element indicates whether the quantum step is in the negative or positive half of the curve. For example, if the first code element were a 1, It would indicate a positive value (e.g., the quantum step is located above the origin). The following three-code elements (bits) identify the segment, as there are seven segments above and seven segments below the origin (horizontal axis).

The first four elements of the fourth + segment are 1101. The first 1 indicates it is above the horizontal axis (e.g., it is positive). The next three element indicate the fourth step or

0-1000 and 1001

1-1010

2-1011

→ 3-1100

4-1101

5-1110 etc

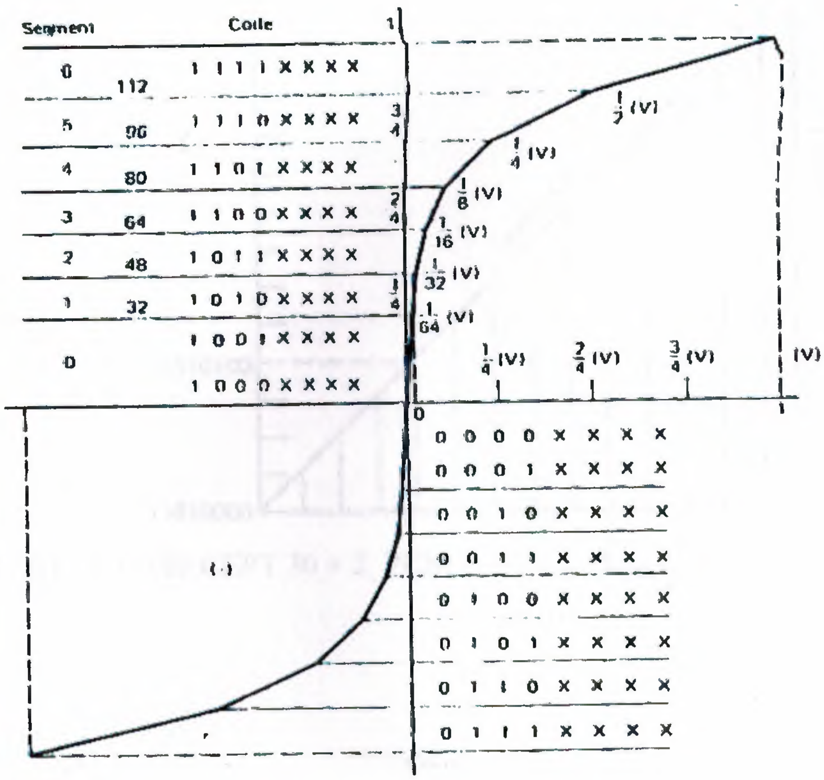


Figure 2.6

Figure 2.7 shows a "blowup" of the uniform quantizing and subsequent straightforward binary coding of step 4. This is the final segment coding, the last four bits of a PCM code word for this system. Note the 16 steps in the segment, which are uniform in size.

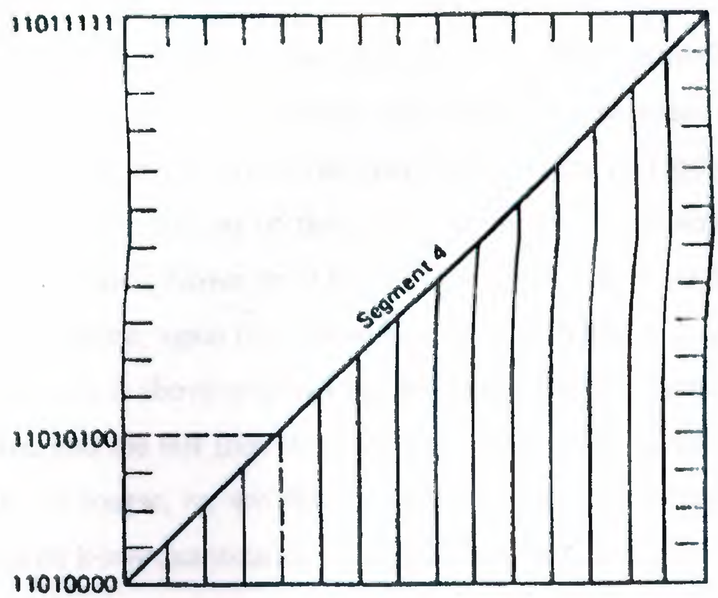


Figure 2.7. The CEPT 30 + 2 PCM system, coding of segment (4 positive).

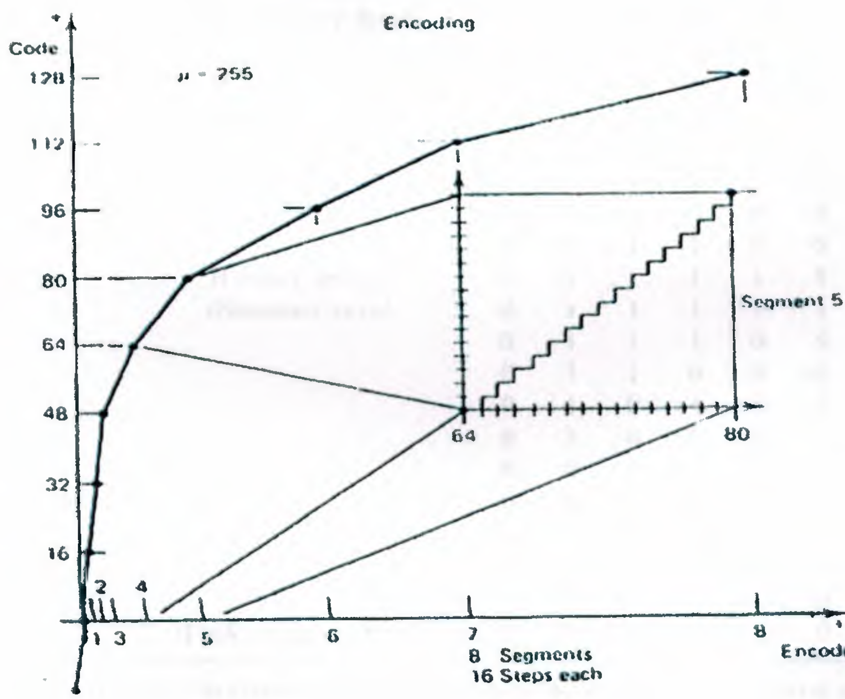


Figure 2.8. Positive portion of segmented approximation of μ law quantizing curve used in North American (ATP) DSJ PCM channelizing equipment. Courtesy of ITT Telecommunications, Raleigh, N.C.

The North American DSI PCM system uses a 15-segment approximation of the logarithmic μ law. Again, there are actually 16 segments. The segments cutting the origin are collinear and counted as one. The quantization in the DSI system is shown In Figure 2.8 for the positive portion of the curve. Segment 5, representing quantizing steps 64 through 80, is shown blown up in Figure 1-8, Figure 1-9 shows the DSI coding. As can be seen in the figure, again the first code element, whether a 1 or a 0, indicates whether the quantum step is above or below the horizontal axis. The next three elements identify the segment, and the last four elements (bits) identify the actual quantum level inside the segment. Of course, we see that the DSI is a basic 24-channel system using eight-level coding with k-law quantization characteristic where $\mu = 255$.

		Digit Number							
Code Level		1	2	3	4	5	6	7	8
255	(Peak positive level)	1	0	0	0	0	0	0	0
239		1	0	0	1	0	0	0	0
223		1	0	1	0	0	0	0	0
207		1	0	1	1	0	0	0	0
191		1	1	0	0	0	0	0	0
175		1	1	0	1	0	0	0	0
159		1	1	1	0	0	0	0	0
143		1	1	1	1	0	0	0	0
127	(Center levels)	1	1	1	1	1	1	1	1
126	(Nominal zero)	0	1	1	1	1	1	1	1
111		0	1	1	1	0	0	0	0
95		0	1	1	0	0	0	0	0
79		0	1	0	1	0	0	0	0
63		0	1	0	0	0	0	0	0
47		0	0	1	1	0	0	0	0
31		0	0	1	0	0	0	0	0
15		0	0	0	1	0	0	0	0
2		0	0	0	0	0	0	1	1
1		0	0	0	0	0	0	1	0
0	(Peak negative level)	0	0	0	0	0	0	1*	0

*One digit is added to ensure that the timing content of the transmitted pattern is maintained.

Figure 2.9. Eight-level coding of North American (ATT) DS1 PCM system. Note that there are actually only 255 quantizing steps because steps 0 and 1 use the same bit sequence, thus avoiding a code sequence with no transitions (i.e., 0's only).

As we know, pulses with two levels, i.e. binary pulses, are attractive for transmission as they are easy to regenerate on the transmission line. It is not difficult to build regenerator circuits able to determine whether a pulse is present or not.

Present-day practical system use binary encoding of the quantized speech samples. See figure 10. As telephony uses 256 quantizing levels, each sample will be encoded to a code group, or PCM word, consisting of 8 binary pulses (8 bits).

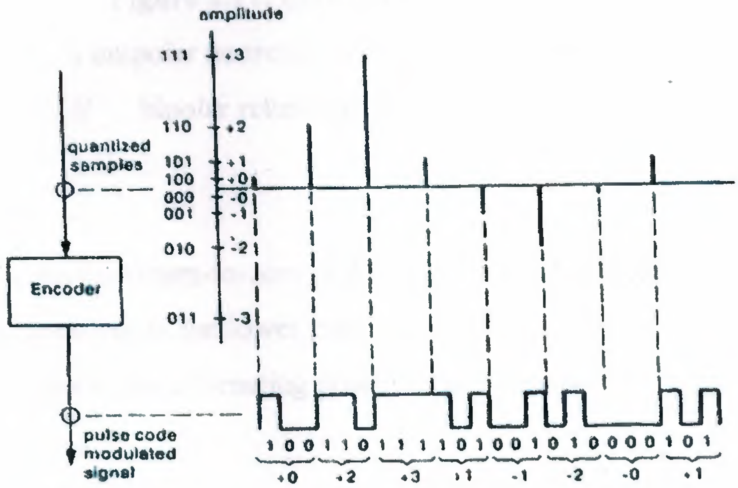


Figure 2.10. Encoding of quantized samples with 8 quantizing levels (3 binary digits/code word).

As the sampling rate used is 8000 samples/second, one pulse code modulated speech signal will generate a 64 kbit/s digital signal.

2.4. Electrical Representation of PCM signals

Digital signals within the terminal are usually transmitted in the form of a unipolar pulse train in the nonreturn-to-zero (NRZ) mode, see figure 11. This signal form is not appropriate for transmission over long distances.

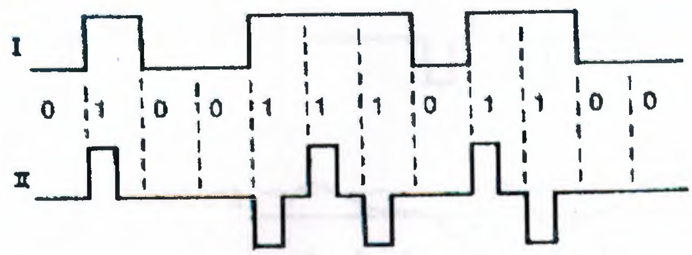


Figure 2.11. Binary information represented in:

I a unipolar nonreturn-1to-zero (NRZ) pulse train.

II bipolar return-to-zero (RZ) pulse trains.

A better form is a bipolar return-to-zero (RZ) signal. The advantages of this signal are

* it has no power in the lower parts of its spectrum, i.e. it has no direct current component; this is due to the alternating polarities of the pulses

* the intersymbol interference is reduced by the return-to-zero feature.

Of course, even this signal will be attenuated and distorted during transmission, and noise will be added to it.

At some point on the transmission line, the signal must be restored. This is done by inserting a device on the line that first examines the distorted pulse train to see whether the likely binary value is 1 or 0, and then generates and transmits to the line new pulses according to the result of the examination. Such a device is called a regenerative repeater. See figure 2.12. At the same time as the pulses are reshaped, the noise added during transmission is eliminated at the least if the noise signal amplitude is not large enough to bring the received code signal to the wrong side of a regenerator decision level. Normally, the regenerated code signal is identical to the transmitted original code signal. Even after a large number of regenerative repeaters. The code signal is practically identical to the original signal. This is the reason for the high transmission quality that is obtainable with PCM transmission system.

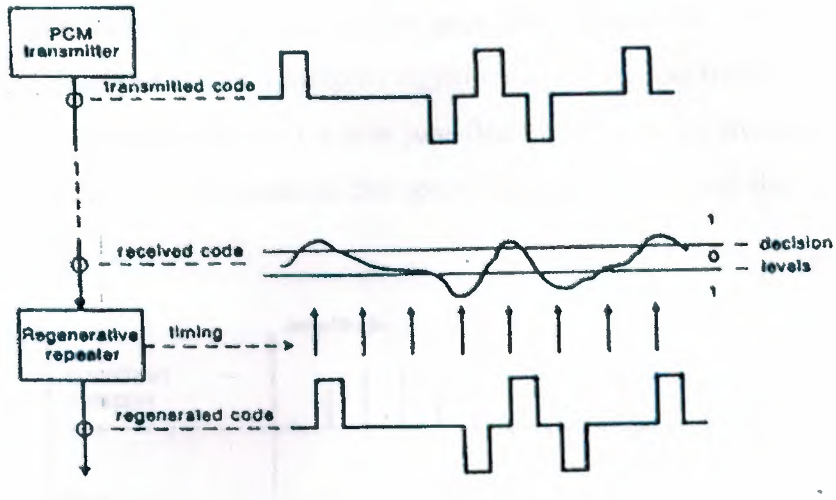


Figure 2.12. PLL15e forms on a transmission line.

2.5. Demodulation

The processes in the receiver that convert the incoming PCM signal to an analogue speech signal again are regeneration, decoding and reconstruction.

The regeneration process has the same aim and is performed in the same way as on the transmission line, i.e. the distorted pulses are replaced by new square pulses, see figure 12. Before entering the decoder, the bipolar signal is reconverted to unipolar. In the decoding process the code words generate amplitude pulses, whose heights are the same as the heights of the quantized samples, which generated the code words. Therefore, after passing through the decoder the train of quantized samples is retrieved. See figure 2.13.

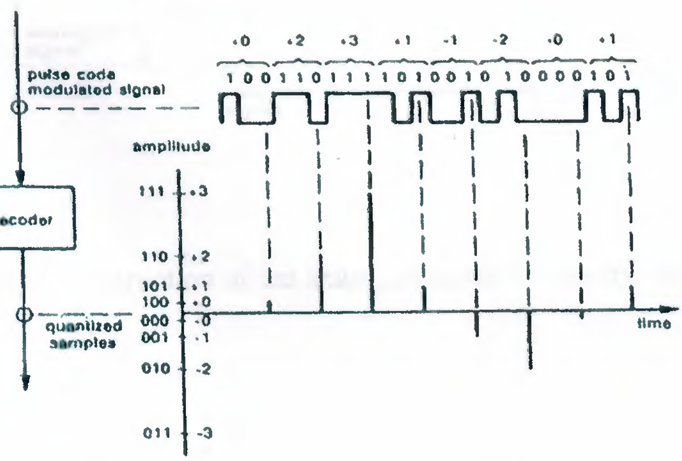


Figure 2.13. Decoding of encoded amplitude levels

The analogue signal is reconstructed in a low pass filter, figure 14a. This can be seen from figure 14b. The spectrum of a sampled signal contains the spectrum of the original signal as has been shown in figure I-3.A low pass filter with a cut-off frequency at B Hz takes away all frequency components in the spectrum above B Hz and the spectrum of the desired analogue signal is left.

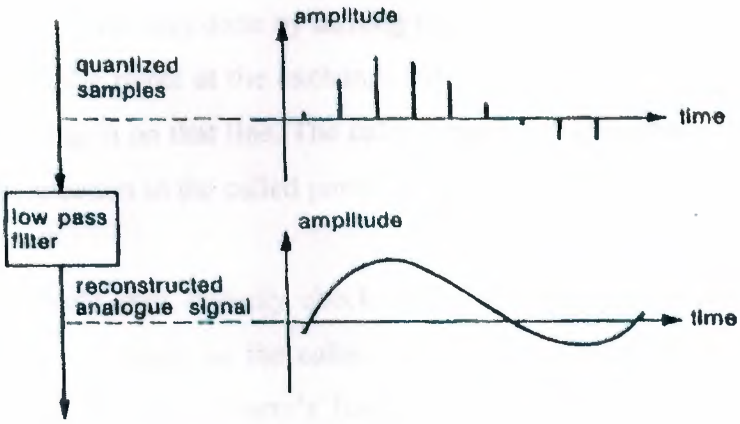


Figure 2.14 a Reconstruction of the analogue

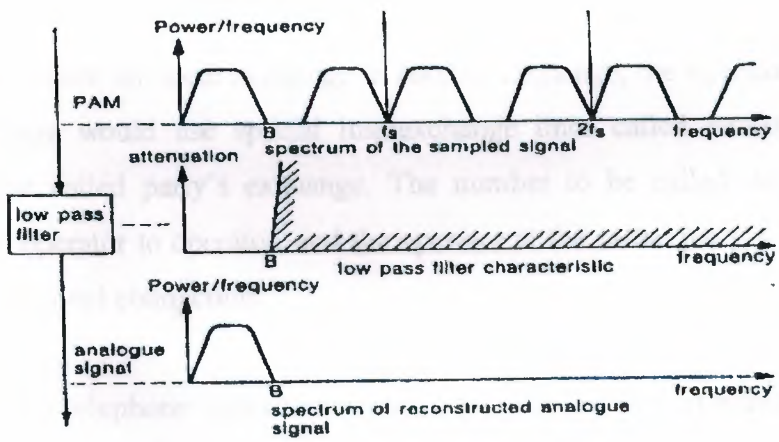


Figure 2.14 b Reconstruction of the analogue signal shown by spectrum diagram

Chapter 3

Working of Telephone Network

3 Signaling

In the early days of telephony, exchange service was accomplished with manual switching by a human operator. The telephone subscriber desiring service first had to alert the operator. This was done by turning the crank on the telephone, which caused a lamp to flash on the panel at the exchange office. The operator would see the flashing lamp and then plug in on that line. The calling party would verbally request the operator to make the connection to the called party.

The operator would then visually check the cords and jacks to determine whether a connection could be made to the called party. If not, the operator would inform the calling party that the called party's line was in use. If the called party's line were available, the operator would make a connection and ring the called party. The lamps for both the called party's line and the calling party's line would remain lit as long as the telephones were in use. As soon as one telephone was hung up, the corresponding lamp would go out, and the operator, noticing this, would unplug the connection.

If the call were from the local exchange to another exchange, the operator at the calling party's exchange would use special interexchange lines called *trunks* to reach the operator at the called party's exchange. The number to be called would be passed verbally from operator to operator, and the operator at the called party's local exchange would make the final connection.

The making of a telephone connection involved a large amount of human labor during the early days of telephony. Technology has, over the years, reduced and finally eliminated all human labor required in making a telephone connection. This was accomplished through automated switching machines and various electrical signals to request service, forward telephone numbers, and set up the actual connection of the lines. The general topic that deals with the various signals used to request service and to control the progress of the telephone call is known as *signaling.*

3.1 Signal Functions

There are four general functions of signals encountered in modern telephony:

- alerting,
- transmitting address information,
- supervising,
- transmitting information signals.

Alerting deals with the initial request for service from the subscriber. In this case, the subscriber sends a signal to the local switching system requesting service. The local switching system might then send signals to other switching systems to alert them that service is requested in the form of interoffice lines, or trunks. A local switching system will then alert the called party to answer the telephone. The telephone number, or address, of the called party must be transmitted from the subscriber to the local switching machine. This is accomplished through either dial pulses or tones. Each switching machine passes the address to the next switching machine. These are all examples of the transmission of address information. Switching machines need to know whether circuits are idle or in use. These machines must also know when a seized circuit is no longer needed and can be released for re-use. The status of circuits thus needs to be supervised.

Information signals must be transmitted to the called party. Such signals as busy tone, dial tone, and various recorded announcements are examples of information signals transmitted to the calling party.

3.1.1 Subscriber Loop Signling

Two major realms of signaling are subscriber-loop signaling and interoffice signaling.

Subscriber-loop signaling involves four functions: alerting, supervision, transmitting address information, and transmitting information signals. Alerting or requesting service is accomplished when the telephone goes off the hook, thereby causing a direct current to flow, which is sensed by equipment at the central office. As long as this direct current flows, the connection is maintained. When this flow of direct current ceases because the telephone is on the hook, the telephone connection ceases. This latter function is supervisory in nature. Thus, dc signaling on the subscriber loop is used for alerting and supervision functions.

Address information can be transmitted in two ways on the subscriber loop. The flow of direct current can be interrupted by the telephone dial to generate dial pulses. These pulses are at a rate of about 10 pulses per second. The second way that address information can be transmitted is in the form of unique two-tone combinations called touch-tone dialing. Information is transmitted on the subscriber loop as either audible tones or recorded announcements. Four major tones are dial tone, ring-back tone, line busy tone, and trunk-busy tone. Four generic frequencies (350, 440, 480, and 620 Hz) are used, either singly, or in combination with each other. Distinctive timing patterns are also used.

Dial tone is a continuous tone formed by combining a 350 Hz sine wave with a 440 Hz sine wave by addition of the two waves. Ring-back or audible ringing is formed by the addition of a 440 Hz sine wave to a 480 Hz sine wave. The combination is on for two seconds and off for four seconds. The line-busy tone is a combination of 480 Hz and 620 Hz sine waves, with the combination on for 0.5 seconds and off for 0.5 seconds. The trunk-busy tone is formed by the same sine waves as the line-busy tone, but the tone is repeated at the faster rate of 0.25 seconds on and 0.25 seconds off.

There is one last alerting signal that is transmitted over the subscriber loop. This is the ringing signal which causes the called telephone to ring. It is a sine wave of 75 volts rms at a frequency of 20 Hz.

3.1.2 Interoffice Signaling

The oldest and most basic type of signaling between central offices is direct current, or dc, signaling. The presence or absence of a dc signal **on** a trunk would indicate whether the trunk was idle or in use. Normal or reverse direct current is sometimes also used for signaling between offices on a per-trunk basis.

Direct current cannot be transmitted over circuits derived from a carrier system, and hence dc signaling cannot be used. The solution was to use a single-frequency tone, either in the voice band (200-3400 Hz), or outside the voice band (3700-3825 Hz). One popular scheme was the use of a 2600 Hz single-frequency tone to indicate whether a trunk was idle or in use.

Unfortunately, there were problems with this in-band, single-frequency, signaling scheme. For one, some people discovered they could generate their own tones for the fraudulent purpose of avoiding toll charges. For another, some speech signals could cause accidental disconnections. The use of an out-of-band frequency was not without its problems, too. One problem was the loss of usable bandwidth for the speech signal. The address information was transmitted over the seized trunk by using two frequency tones sent at a rate of 10 tones per second. The tones consisted of two-frequency combinations of 700, 900, 1100, 1300, 1500, and 1700Hz. With the various combinations of these frequencies, it was possible to represent all ten digits and up to six control functions. This type of ac signaling was called multiple frequency key pulsing (MFKP).

3.1.3 Common Channel Interoffice Signaling

In 1976, a new interoffice signaling scheme was first installed in the Bell System, and it is currently in use on practically all interoffice trunk circuits. This new system is called common channel interoffice signaling or CC1S for short.

With CCIS, as shown by Figure 3-1, a separate channel between the offices is dedicated as a data link for transmitting only signaling information. No signaling information is sent over the voice circuits. The switching machines used in most central offices are actually computers or processors that control the switching of voice circuits or paths. As such, the use of a data link to enable these computers to communicate with each other about the availability of the voice circuits between offices is quite consistent with the capabilities and requirements of the newer switching technology.

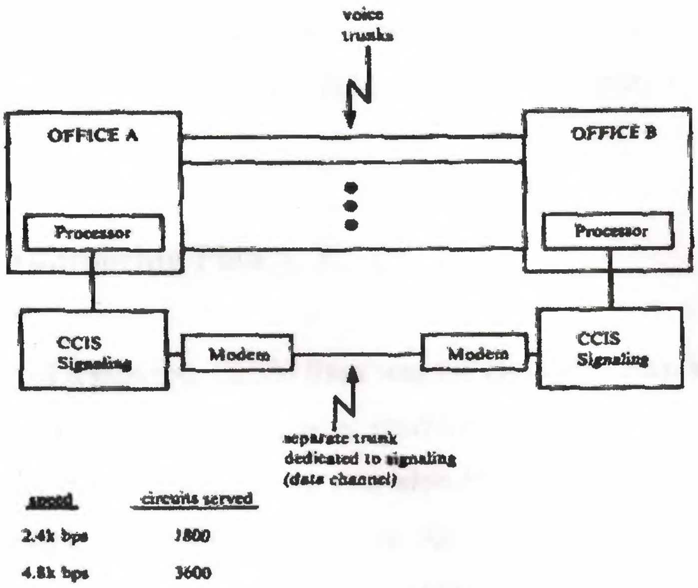


Figure 3.1

A single analog circuit is used to convey the digital signaling information. Conventional full-duplex modems operating at either 2.4 k bps or 4.8 k bps are used. Each signaling circuit can control about 1800 or 3600 voice circuits, respectively.

Because the signaling information is transmitted over a separate circuit, there is a need to determine whether the transmission quality of the specific voice circuit is acceptable before it is connected for use. This is accomplished by performing a transmission quality check on each voice circuit before it is connected for service. The voice circuit is looped back onto itself, and a tone is transmitted down the circuit. The return level of the tone is checked to be certain that it is within specifications. With conventional signaling, a busy tone is sent from the office closest to the called party, all the way back down the network to the calling party, thus tying up a full voice circuit. With CCIS, a busy tone is generated at the office closest to the calling party, hence freeing voice circuits for use with actual conversations.

CCIS at present does not transmit the calling party's identification to the terminating office. However, CCIS could be given the capability of doing so, which would make possible call screening at the local terminating central office. Also, it is possible to envision that the calling party's identification might be transmitted all the way down the local loop so that the called party would know the identity of the calling party before answering the telephone. CCIS, therefore, might make possible many new services in the future.

3.1.4 Telephone Numbering Plan

During the early days of telephony, 10,000 lines was the maximum number served by a telephone exchange. Thus, a four-digit number specified the party to be reached in an exchange. The exchange was specified by two alphabetic characters followed by a decimal digit, for example, WA5 for Waverly-five. Area codes were then introduced to specify the area in the country to be reached. Area codes are also called numbering plan areas (NPA).

A special nomenclature is used to describe the telephone numbering plan. The symbol *N* is used for any of the decimal digits 2 through 9; the symbol X for any of the decimal digits 0 through 9; and 0/1 for the digits 0 or I only.

The standard format for telephone numbers in the United States has been NO/I *X-NXX-XXXX.* N0/I *X* specified the NPA; *NXX* gave the local exchange in the NPA; and XXXX denoted the specific subscriber line in the local exchange. Because the number of area codes possible with the N0/ *IX* format is being gradually exhausted; a new format of *NXX* is being introduced for the NPA.

The format *N* 11 is used for special services. For example, 411 specifies directory assistance; 611 is the repair service; and 911 is for emergencies.

3.1.5 Local Loop Signaling Design

The resistance of the local loop must not be too high, otherwise not enough current will flow in the line to activate the line relay at the central office. The resistance of the local loop depends on the total length of the loop and the gauge of the wire. Resistances for various gauges of wire are as follows:

26 gauge — 83 ohms per 1000 ft,

24 gauge — 53 ohms per 1000 ft.

22 gauge — 32 ohms per 1000 ft.

19 gauge — 17 ohms per 1000 ft.

The maximum resistance of the local loop can be calculated as follows. The telephone instrument requires about 23 mA for the carbon granule transmitter to operate reliably. The common battery at the central office has an electromotive force of 48 volts. Thus, the total resistance of the circuit must not exceed 48/0.023 2100 ohms. The resistance of the telephone instrument is equivalent to 400 ohms. Similarly, the resistance of the central office circuitry is also 400 ohms. Hence, the resistance of the loop must not exceed 2100 - 800 1300 ohms.

The maximum loop resistance of 1300 ohms determines the wire gauge for a given loop length.

3.2 Transmission

The aspect of telephony that deals with the various media and technologies for conveying the telephone speech signal from one place to another is called transmission.

There are a variety of different transmission media that are used to transmit telephone speech signals. Some of these media can carry only a single speech signal, while others can carry many speech signals multiplexed together through either frequency-division multiplexing (FDM) or time-division multiplexing (TDM). The various media are as follows:

- open wire,
- paired cable (commonly called twisted-pair),
- coaxial cable,
- microwave radio (terrestrial and satellite paths),
- optical fiber.

These media vary greatly in terms of the number of speech circuits that they can carry.

3.3 Open Wire and Paired Cable

Open wire consists of pairs of uninsulated wires that are strung on poles. The wires in each pair are physically separated by a distance of about one foot to prevent short circuits during high winds. Open wire has a low loss, typically about 0.03 dB per mile, and was used during the early days of telephony until physical congestion became a serious problem. Open wire is still found, although infrequently, in rural areas.

A twisted pair is a pair of insulated wires twisted together with a full twist about every 2 to 6 inches. The insulation is typically plastic, but wood pulp has been used in the past. The diameter of the wire varies from 0.0 16 inches (26 gauges) to 0.036 inch (19 gauges). Many twisted pairs are combined together into a single cable, usually sheathed with plastic, although older cables were sheathed with lead. Anywhere from 6 to 2700 twisted pairs are combined together into a single paired cable. The gauge of the wire varies with the number of twisted pairs in the cable; finer wire is used in larger capacity cables. Paired cable can be strung on poles, buried underground, or installed in a conduit. A conduit consists of large blocks of concrete with holes through which the cable passes. Conduit is buried underground. A conduit offers the advantage that cable can be replaced, without digging up city streets, simply by pulling out the old cable and pulling through the new cable.

The thinner wire used in paired cable has a higher loss than open wire. The heaviest gauge wire (19 gauge) has a loss of about 1.1 dB per mile at 1000 Hz. Popular paired cables contain 2700 pairs of 26-gauge wire, 1800 pairs of 24-gauge wire, and 110 pairs of 22-gauge wire.

A problem with many wire pairs which are all running parallel to each other with close spacing is that the electrical signal on one pair can leak to another pair. This effect is called crosstalk. Paired cable is mostly used for the local loop, and also between local-exchange central offices. Baseband transmission is used on most local loops, but in cases of severe congestion, subscriber loop carrier (SLC) systems are available, using either frequency-division multiplexing via amplitude modulation (AM) or time-division multiplexing via digital carrier systems.

Analog frequency-division multiplexing and digital time-division multiplexing are used on paired cables in exchange trunk transmission when congestion necessitates an enhancement of baseband transmission. These multiplexing schemes are used on short-haul (15 to 200 miles) trunks. The analog systems are called N-carrier, and typically multiplex together 12 or 24 voice circuits. The digital systems are called T-carrier, and multiplex together 24 voice circuits.

3.3.1 Loading Coils

An electrical transmission line can be modeled as a series of infinitesimally small elements consisting of a series inductance *L,* a series resistance R, a shunt (or parallel) capacitance *C,* and a shunt resistance S. These various quantities are expressed in ohms, henrys, and farads on a per unit length basis, for example, ohms per mile.
At telephone frequencies, the attenuation *A* (or loss) of such a line is given approximately by the following equation:

$$A = R/2 \sqrt{C/L} + \sqrt{L/C}$$

The first term represents the effects of the series losses, and the second term represents the effects of the shunt losses. Usually, the series losses predominate. Thus, the introduction of additional series inductance will decrease the predominant first term, resulting in a decrease in the overall attenuation of the line. Clearly, if too much series inductance is added, then the second term could become predominant, thereby negating the desired effect of reducing the overall attenuation.

The required series inductance is added as discrete inductors placed in series every 6000 feet along the line. An inductor is a coil of wire, and since these series inductors load the line to reduce attenuation, they are called loading coil.c. Their inductance is about 88 millihenrys.

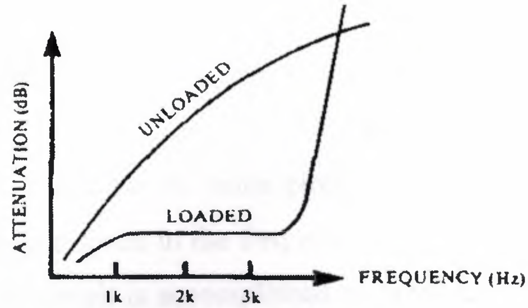


Figure 3.2

Loading coils were invented in 1899 by both Dr. Michael I. Pupin of Columbia University and AT&T employee George A. Campbell. The patent was awarded to Dr. Pupin based on a disclosure only two weeks earlier than Campbell's. The theoretical analysis on which the inventions were based was performed by Oliver Heaviside in England in the late 1800s. Although a loading coil reduces attenuation in the voice band, attenuation outside this band is greatly increased. The first loading coils were installed experimentally in 1899. They were quickly adopted for use on most long cables, and today are in use on long local loops. The inductance is chosen to ensure a passband from 300 to about 3300 Hz. The type of loading most popularly used is specified as *H*-88. The *H* signifies a 6 kilofeet spacing, and the 88 signifies an inductance of 88 mill henrys.

3.3.2 Multiplexing

A number of voice circuits are all multiplexed together on a single transmission medium as illustrated by Figure 3.3. There are two approaches to multiplexing: analog, or frequency-division multiplexing, and digital or time-division multiplexing.

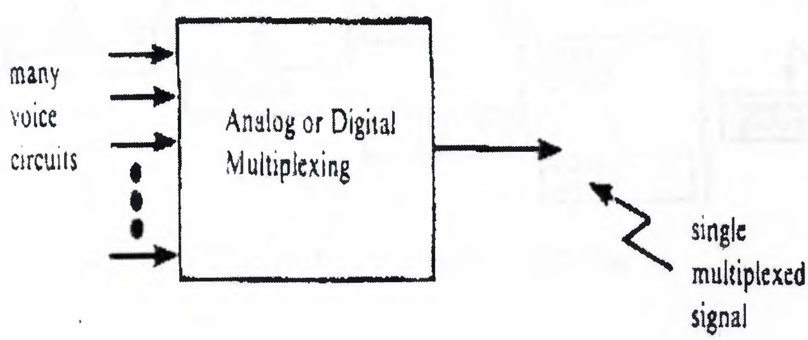


Figure 3.3 Analog or Digital Multiplexer

3.3.3 Analog Multiplexing

In analog multiplexing, a number of voice circuits are combined together, with each circuit given its own unique space in the frequency spectrum. The frequency translation of each baseband speech circuit is accomplished by amplitude modulation using single-sideban, suppressedcarrier modulation. The actual multiplexing is accomplished as a multileveled process in which a small number of circuits is multiplexed together to form groups, and these groups are then multiplexed together, and so forth.

A speech or voice consists of baseband frequencies in the range from 200 to 3400 Hz. A single voice circuit occupies a channel. Twelve channels are multiplexed together, each channel being 4 kHz wide to create a group. A group covers the frequency range from 60 to 108 kHz. Five groups are multiplexed together to create a supergroup occupying the frequency range from 312 to 552 kHz.

This process can be continued (Figure 3.4). Ten supergroups multiplexed together give a mastergroup occupying the frequency range from 564 to 3084 kHz and_containing a total of 600 voice channels. Six mastergroups multiplexed together give a jumbogroup occupying the frequency range from 564 to 17,548 kHz and containing 3600 channels. Three jumbogroup multiplexed together give a jumbogroup multiplex containing 10,800 channels.

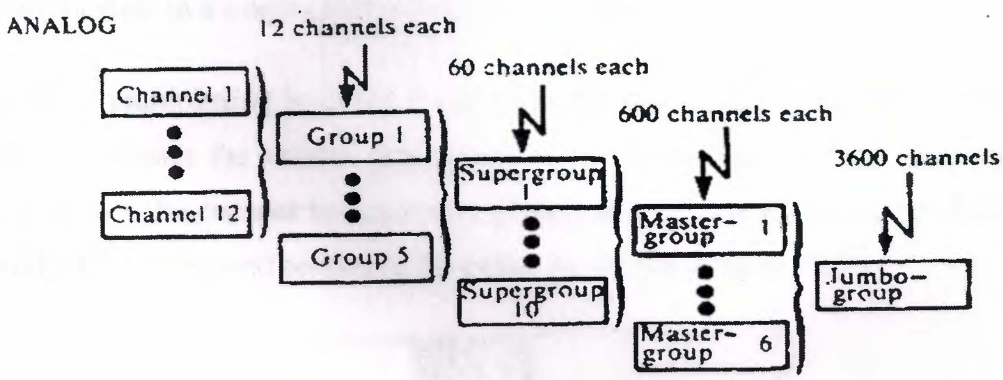


Figure 3.4 Analog Multiplexing

3.3.4 Digital Multiplexing

The equipment that performs the multiplexing is called a channel bank. Analog or A-type channel banks perform analog multiplexing, and digital or D-type channel banks perform digital multiplexing.

Twenty-four voice channels digitally multiplexed together give a DS-l signal, requiring a data rate of 1.544 million bits per second (M bps). A group of 24 voice channels digitally multiplexed together is sometimes called a digital group or a *"digroup"* for short. Four DS- 1 signals digitally multiplexed together give a DS-1 containing 96 channels and requiring a data rate of 6.312 M bps. Seven DS-2 signals digitally multiplexed together give a DS-3 signal, containing 672 channels and requiring 44.736 M bps. Six DS-3 signals digitally multiplexed together give a DS-4 signal, containing 4032 channels and requiring a data rate of 274.176 M bps.

3.4 Coaxial Cable

Analog multiplexing on coaxial cable has been used in the Bell System since 1946 for long-distance telephone transmission. A number of one-way voice circuits are frequency multiplexed together using single-sideband, suppressed-carrier amplitude modulation on a single coaxial cable. Two such coaxial cables make a two-way pair, with each cable carrying transmission in one direction. A number of coaxial-cable pairs are placed together in a single cable to make a total transmission system.

Coaxial cable multiplexing is called L-carrier in the Bell System. A number suffixed after the L. indicates the various generations of the technology. A key factor in L-carrier systems is the distance between the repeaters at which the signal is amplified and retransmitted down the next section of the cable. As shown in figure 3.4.a

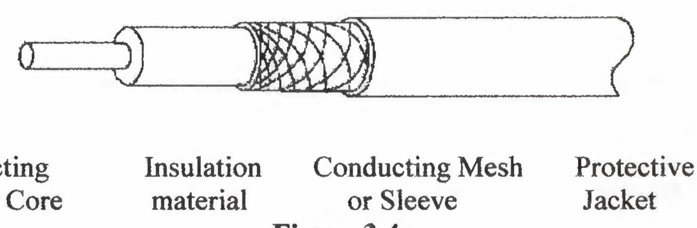


Figure 3.4a

The table 3.1 shows the progression over time of L-carrier system. The most recent system is L5,which was first in service in 1974. The L5-carrier system used integrated circuit technology with repeaters spaced every mile along the route. As with all L-carrier systems, the cable is buried underground.

	L1	L3	L4	L5
● SERVICE DATE	1946	1953	1967	1974
● TECHNOLOGY	vacuum tubes	vacuum tubes	transistors	*integrated* circuits
● REPEATER SPACING (miles)	8	4	2	1
● CAPACITY PER COAX *(channels)* (groups)	600 master group	1860 3 master and 1 supergroup	3600 jumbo group	10,800 jumbogroup multiplex
● COAX PAIRS	4	6	10	11
● WORKING PAIRS	3	5	9	10
● ROUTE CAPACITY *(two-way voice circuits)*	1800	9300	32,400	108,000

Table 3.1

The channel capacity of a single coaxial cable in L5-carrier is a jumbogroup multiplex (10,800 channels). There are 11 coaxial cable pairs, 10 of which are in actual service. Thus, the overall route capacity of L5 is 108,000 two-way voice circuits.

3.5 Micro Wave Terrestrial System

Radio transmission is used to carry telephone conversations across continents and oceans. Different frequency bands have been allocated for use in telephone radio transmission by the common carriers. A large number of voice circuits are multiplexed together in these radio systems. The microwave radio bands are in the gigahertz (10^9)

range of frequencies, and are used in cross-country terrestrial and satellite routes. Two bands currently used are from 3.7 to 4.2 GHz and from 5.925 to 6.425 GHz. The width of each of these bands is 500 MHz. The radio channel width in the first band is 20 MHz and 3O MHz in the second band. The first band is called the 6GHz band, and the second band is called the 6 6Hz band. These two bands are used for microwave terrestrial radio. The extremely high frequencies of microwave radio are conducted through metal pipes called waveguides before being transmitted over the air at the antenna.

Microwave terrestrial radio forms the bulk of the long-distance telephone network. The microwave radio beam follows a line-of-sight path which necessitates that a series of towers be located about every 26 miles, on the average across the route of the system, an antenna on each tower receives the radio signal, and the signal is amplified and rebroadcast to the next tower. The towers perform as repeater stations. For the 4 GHz and 6GHz bands, rain does not usually have significant effect on the propagation of the radio wave. However, rain does have a significant effect on the 11 GHz (10.7 to 11.7 6Hz) band and the 18 6Hz (17.7 to 19.7 GHz) band, which are also used for microwave transmission.

3.6 Communication Satellites

Terrestrial microwave radio is not suitable for transmission across oceans because its line-of-sight nature would require towers every 26 miles across the water. The solution is a single microwave tower placed high enough in the sky that the whole distance could be covered in a single hop. Since neither sky-hooks nor towers hundreds of miles high are feasible, some practical form of implementation is needed.

The Bell Labs system would have consisted of a large number of satellites at the relatively low altitude of about 3000 miles. At this altitude, satellites are not stationary with respect to the earth's rotation, and hence a series of satellites would need to be continuously tracked across the sky as they passed overhead. There were serious questions about whether such a system was technically and economically practical. The final solution was a single satellite at a height such that the orbit time of the satellite was the same as that of the earth, so that the satellite would appear stationary with respect to the surface of the earth.

The curvature of the earth's surface will be falling away at the same rate as the shot is being pulled back by gravity toward the surface. The shot will then continuously fall around the earth! It will have become an artificial satellite of the earth. The height of the orbit of a satellite and its rate of rotation about the earth are related. The higher the orbit, the slower is the rate of rotation about the earth. There is a height at which the time to complete one orbit is actually the same time it takes the earth to complete one full rotation. If the orbit of the satellite is exactly above the earth's equator and in the same direction that the earth is turning, the satellite will appear stationary with respect to the earth's surface. This type of orbit is called a geosynchronous orbit, and the height is 22,300 miles above the earth's equator, or equivalently 26,300 miles from the earth's center. (See Figure 4-15.) The use of geosynchronous communication satellites (also called geostationary) was first suggested in 1945 by Arthur C. Clarke.

The electronic circuitry on a satellite receives the signal transmitted to it from the earth station. This signal is very weak and must be amplified by low-noise amplifiers on board the satellite. The signal is then changed in frequency and retransmitted back to earth. These operations are performed by circuitry called a transponder. Each radio channel has its own transponder, and thus a number of transponders are on board the satellite to cover the whole frequency band allocated to it. Modern communication satellites typically have 24 transponders.

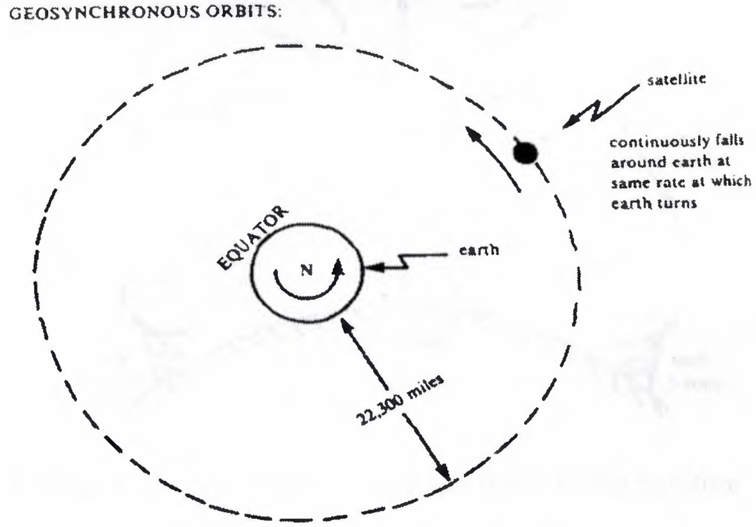


Figure 3.5 Geosynchronous Orbit

3.6.1 Frequency Band

Different frequency bands must be used for transmissions from the earth to the satellite and from the satellite to the earth. (See Figure 3.6 a.) The different frequency bands are required to eliminate interference between the received and transmitted signals. The band of frequencies used to transmit from the earth to the satellite is called the up-link, and the band of frequencies used to transmit from the satellite to the earth is called the down-link.

The earliest frequency bands used for satellite transmission were the same 4 GHz and 6 GHz bands used for terrestrial microwave transmission. The 4 GHz (3.7 to 4.2 GHz) band was used for the down-link, and the 6 GHz (5.925 to 6.425 GHz) band was used for the up-link. The 4 GHz and 6 GHz bands used together for satellite transmission is called the C band (see Figure 3.6 b). Newer communication satellites are beginning to use the Ku band at 12 GHz and 14 GHz. International agreements have authorized the use of a third band, the Ka band, at 17 0Hz and 30 GHz.

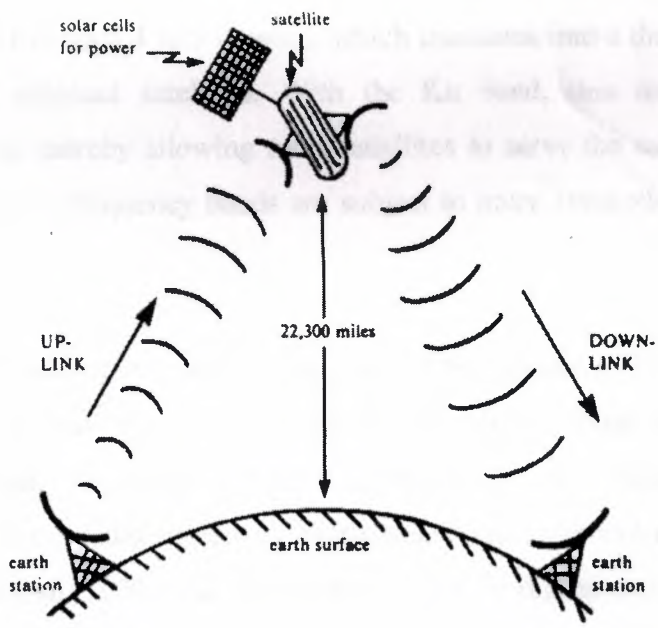


Figure 3.6.a Transmission from the Earth to the Satellite

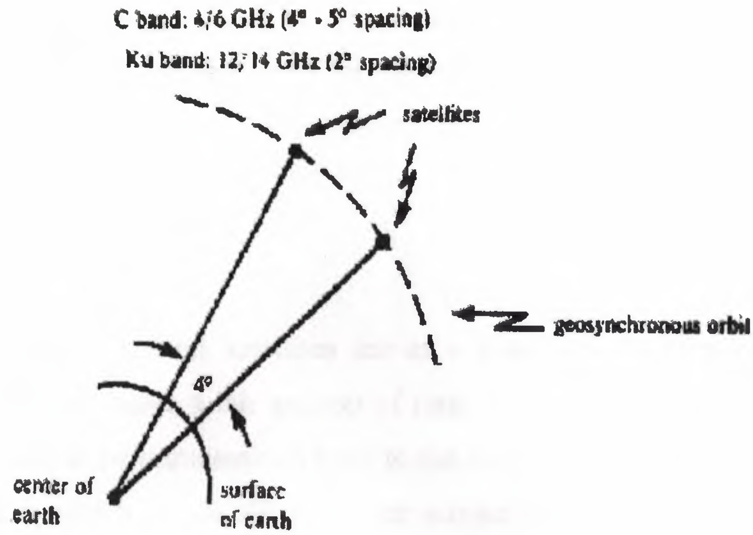


Figure 3.6 b

The higher frequency Ku and Ka bands allow the transmitted radio beam to be very narrow. Because all satellites in the same geosynchronous orbit use the same bands, it is essential that the signal sent to one satellite not be received by an adjacent satellite and, similarly, the earth antenna aimed at one satellite should not receive the signal from an adjacent satellite. These requirements and the width of the radio beam determine the permissible spacing between the satellites in geosynchronous orbit. The permissible spacing in the C band has been 4 to 5 degrees, which translates into a distance of about 2000 miles between adjacent satellites. With the Ku band, this distance can be decreased to 2 degrees, thereby allowing more satellites to serve the same geographic area. However, the higher frequency bands are subject to more absorption of the radio signals by rain.

The typical spectrum width of the radio channel served by one transponder is 36 M Hz. Because reception of a weak signal in the presence of noise is prime consideration in satellite communication, the noise immunity of frequency modulation is used for modulating the microwave radio signals. A single transponder channel can be used for one color television signal, 1200 voice telephone circuits, or digital data at a rate of 50 M bps.

The total frequency bandwidth of the C band is 500 MHz. Vertical and horizontal polarization of the radio signal is used to achieve a total capacity of 24 channels, each 36 M Hz wide. A communication satellite allocates these channels to create 12 two-way transponder pairs.

3.6.2 Delay

Geosynchronous communication satellites are at a considerable distance above the surface of the earth, and hence a fair amount of time is required for the radio signal to reach the satellite and to be retransmitted back to the earth. Radio waves travel at nearly the velocity of light, which is 186,000 miles per second. A radio wave will, therefore, take about 120 milliseconds (ms) to travel the 22,300 miles to the satellite. Actually, the earth station is not directly below the satellite, and hence a little extra time needs to be added for this extra distance. A net time of 135 ms is reasonable for the radio wave to travel from the earth station to the satellite.

The signal received by the satellite is retransmitted back to earth. An additional 135 ms is required for the radio signal to travel that distance. Thus, there is a total delay of about 270 ms for a signal to be sent by satellite from one location on the surface of the earth to another, assuming that a single satellite hop is used.

For telephone communication, one person speaks and the other person responds. It would take 270 ms for the speech to travel from one person to the other, and the additional 270 ms for the response to travel back via a return satellite circuit. The total round-trip delay would be about 540 ms, or about a half second. This amount of delay is quite noticeable to most people and would be quite bothersome to many. If one were using a terminal to a distant computer via a satellite link, this round-trip delay would add appreciably to the response time. Also, this delay would be intolerable for the situation in which one computer is communicating interactively with another.

In some communication situations, one satellite hop might be connected to a second satellite hop. The delay in these multihop satellite links can bring interpersonal telecommunication to a virtual halt, unless each person learns to wait patiently to receive the other person's response.

A possible solution to multihop links would be to beam signals between the satellites themselves, thereby eliminating the need for intermediate transmission to the earth. The use of lasers would be particularly applicable to this type of space communication because there is no optical interference in the vacuum of space.

3.6.3 Access

The size and shape of the satellite's radio beam on the surface of the earth is called the footprint of the satellite. Clearly, for a satellite that is broadcasting signals to as wide an audience as possible, the footprint should be large. If the satellite is transmitting signals to only a specific earth station, then the footprint should be small. In the first case, the radio beam must be broad, and in the second case the radio beam would need to be narrowly focused. The actual antenna design and available transmitting power on the satellite are important factors in determining footprints.

It is possible to design the satellite's antenna so that it can send narrowly focused beams to a number of specific, geographically dispersed locations. It is also possible to switch the narrowly focused beam from one location to another. The specific application for the latter would be to transmit information to one location and then to another. If broad beams are used, then everyone receives the same information, and security may also become a consideration.

In the same way that one satellite beams information to a number of users at different locations, a number of locations may also beam information to the satellite. One way in which this can be accomplished is for each multiple user to be allocated a specific band of frequencies for each transmission. This
type of access is called frequency-division multiple access(FD MA). Another n way to accomplish multiple access is to give time slots to each multiple user, .a during which time users must transmit their information. This type of access is called *time-division* multiple access (TDMA). TDMA is particularly well suited to the bursty nature of data communication.

In addition to the actual communication signals carried by satellites, telemetry signals are transmitted to and from the satellite. These signals report on the state of the satellite with respect to such parameters as the temperature, battery life, and orbital position. If the position of the satellite needs to be corrected because of drift, telemetry signals from earth will activate small thrusters on the satellite to reposition it. Satellites are powered by solar energy using rays of light from the sun. This light energy is converted into electrical energy by panels of solar cells on board the satellite. The efficiency of this solar conversion and the physical size of the panels determine the total electrical power on the satellite, in addition to the power available to transmit signals back to the earth. For a portion of each orbit, the satellite will be shielded from the sun by the earth. During this time, batteries that were previously charged from the solar panels are used to power the satellite.

3.6.4 Capacity

The communication satellite Corns ear I was launched in 1976. It used 1200 voice circuits per transponder for a total capacity of 54,400 two-way voice circuits. Vertical and horizontal polarization of the radio waves was used for the 24 transponders operating in the 6/4 GHz C band. Comstar IV was launched in 1980. Improvements in technology allowed 1800 voice circuits to be used for each transponder channel, thereby giving a total capacity of 21,600 two-way voice circuits. Both Comstar I and Comstar IV used frequency modulation of the radio wave. In 1982, the equipment at the earth stations was changed so that single-sideband amplitude modulation (SSB/ AM) could be used. The number of circuits that could be handled by each transponder channel was thus increased to 7800 for a total capacity of 93,600 two-way voice circuits.

Future satellites will continue to use SSB/ AM, probably increasing the number of circuits per channel to 12,000 voice circuits. Overall capacity will thus be increased proportionally.

3.7 Optical Fiber

An optical boundary is formed when two materials having different indices of refraction are joined. A light wave passing across this boundary is bent when it emerges in the other material. If the light wave impinges on the boundary at a very shallow angle, called the critical angle, it will be reflected at the boundary, as if the boundary were a perfect mirror.

Imagine a long mirror that has been rolled up to create a long tube, with the mirrored surface on the inside of the tube. If a source of light is shined into one end of the tube, the light waves will be repeatedly reflected off the surface of the mirror and guided along the length of the tube until they emerge at the end. In effect, the mirrored tube is a light, or optical, waveguide. Consider a long, thin, fiber core made of glass, surrounded by a cladding of glass with a lower index of refraction. If a light wave enters the core at a shallow angle, it will be repeatedly reflected at the boundary of the core and the cladding until it emerges from the end. This is the basic principle of the optical fiber. It is a medium for transmitting light waves.

The use of glass pipes and fibers to guide light is an old idea. American William Wheeler was granted a patent in 1881 on the use of mirrored glass pipes for guiding light, and the use of thin glass fibers for guiding light was described in 1887 by the British physicist, Charles Vernon Boys. Alexander Graham Bell invented the photophone in 1880. It transmitted voices over the air on light waves. The synthesis of Bell's ideas with the light-carrying capabilities of thin glass fibers did not occur, however, and it was not until quite recently that optical fibers have come into widespread use for communication. Also, digital multiplexing was not known a century ago, nor was the required electronic circuitry available.

The exterior of an optical fiber is coated with a jacket (see Figure 3.7) to prevent stray light from entering the fiber. The actual glass from which the optical fiber is manufactured has an extremely high purity. Typical losses are from I to 4 dB per km. and even purer glass fibers are being developed as the technology continues to progress.

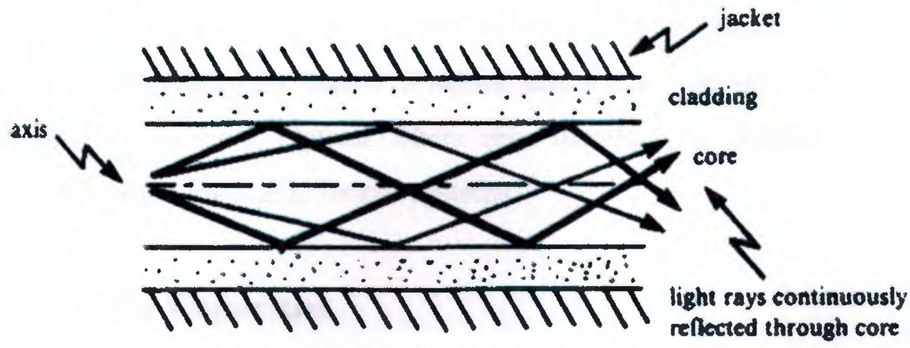


Figure 3.7 Optical fiber

Optical fiber is exciting as a transmission medium because of the very high frequency of the light signal which it transmits. This means that a very large amount of information can be transmitted over a single optical fiber. The maximum amount of information that can be transmitted over an optical fiber is determined by the *smearing* of a pulse that occurs because of the multiple path lengths that the light rays take in traveling through the fiber (see Figure 4-23). Some types of optical fiber are worse than others in terms of smearing.

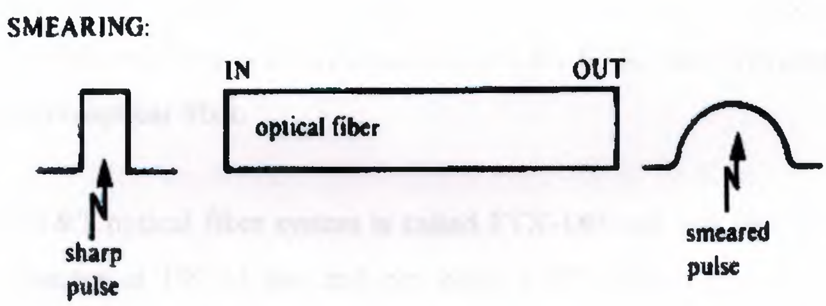


Figure 3.8 Smearing

An optical fiber communication system consists of a light source, the fiber itself, and a light detector. The light is detected and new pulses are generated along the length of the system at regular intervals determined by the degree of smearing, noise, and losses in the fiber. The regeneration of pulses along the length of the system is called repeating and is performed by regenerative repeaters.

Because there is no practical way at present to amplitude or frequency modulate light waves, optical fiber systems are digital in nature, and simply turn the light on and off to convey digital information. Optical fibers are, therefore, a transmission medium particularly well suited for digital carrier systems.

3.8 Transmission Capacities

The various transmission technologies that are currently available for long-distance telephone communication are:

- coaxial cable,
- terrestrial radio,
- satellites,
- optical fiber,
- undersea cable.

The newest technology is optical fiber, and it is already the clear winner in maximum route capacity. The first system, FT3, was introduced in 1979 and used graded-index fiber at 45 M bps, capable of carrying 672 voice circuits (one DS-3 digital channels). The maximum route capacity consisted of 72 fiber pairs, with 66 in service, for a total capacity of 44,352 two-way voice circuits. The AT&T FT3C was previously described in the section on optical fiber.

The newest AT&T optical fiber system is called FTX-180 and was introduced in 1984. Each fiber operates at 180 M bps and can carry 2,688 voice circuits. A 400 M bps system using single-mode fiber is scheduled for 1986 with each fiber carrying 6048 circuits. The repeaters will be spaced every 20 miles in this system. The total route capacity will depend on the number of fiber pairs a single coaxial cable and required 660 repeaters.

The SD transoceanic system utilized vacuum tubes and became avail. able in 1963 with 140 two-way voice circuits on a single coaxial cable. T~ SF system utilized transistors and became available in 1968, offering 84Otwod way circuits on a single coaxial cable. The capacities of these two system could be doubled through the use of TASI.

3.9 Echo Elimination

Echo can be a serious problem on toll circuits, and, therefore, it must be eliminated. One solution to the problem is to break or open the circulating path in the four-wire toll circuit. Devices called echo suppressors accomplish this.

As illustrated by the diagram of Figure 3.9, an echo suppressor inserts loss in the return path so that the echo caused by any leakage across the hybrid, or by reflections on the local loop, cannot return down the four-wire toll circuit. The loss in the return path is activated if any speech signal is sensed at the input to the hybrid. The output from the hybrid is compared with the input to the hybrid to determine whether the near-end party is speaking. If so, then the loss must be removed from the return path. Echo suppressors should be carefully designed so that the switching action occurs quickly, but with optimum dynamics so as to interfere as little as possible with the interactions of an interpersonal conversation. Echo suppressors have problems because two-way talking is not possible. Once the distant speaker has activated the loss, the other person must speak loud enough to deactivate the loss in order to be heard. The echo suppressor has, in effect, made the circuit half-duplex. The use of the proper amount of loss Fl enables the other party to be heard weakly, and thus minimizes the half-duplex effect.

The problems of echo suppressors are avoided with echo cancellers. Echo cancellers create a synthetic echo that is subtracted from the return signal, thereby canceling the echo, according to the schematic in Figure 3.8, use any local loop with varying electrical characteristics can be the obtained, it is impossible to know the echo exactly in advance. The cancellation, therefore, must be performed with complex electronic filters, which the adjust their properties dynamically to minimize the returned signal. This of dynamic filtering is performed by digital filters. Echo cancellers have become on economical by the use of integrated circuits. Their one disadvantage is that a short amount of time is required for them to adapt to the connection and determine the appropriate filter response to minimize the echo. However, in practice, this adaptation time is not noticeable to most people. Of course, two-way talking and full-duplex operation is made possible with echo cancellers.

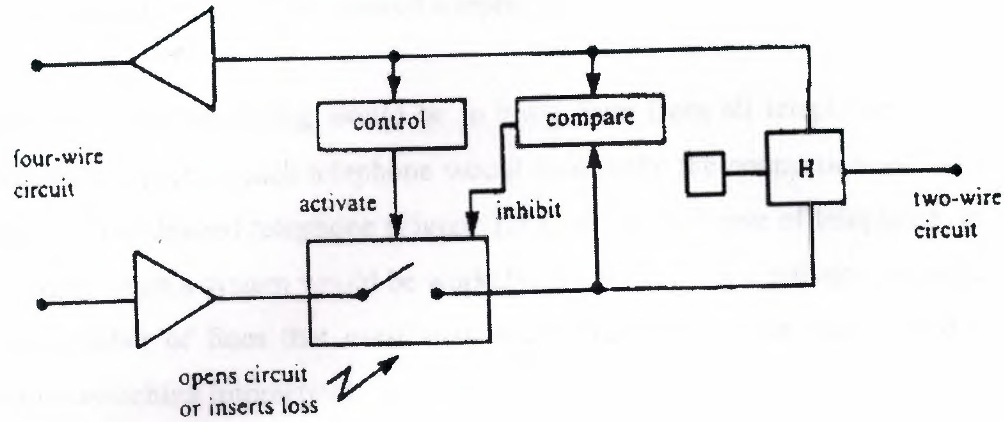


Figure 3.8.a Four-wire Toll Circuit

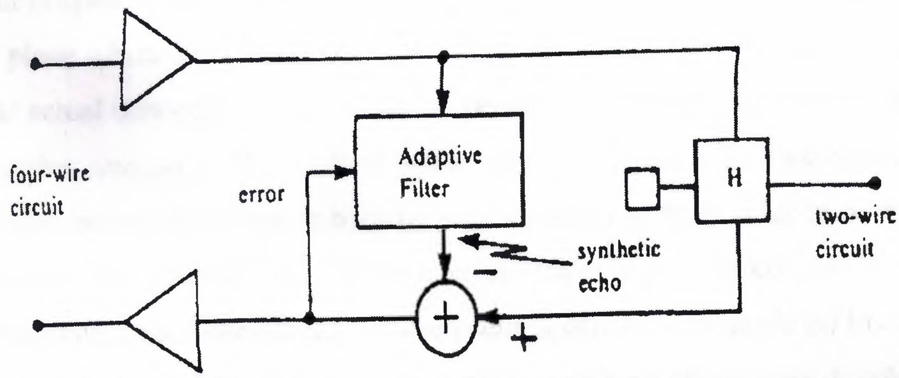


Figure 3.8.b Echo Suppressor

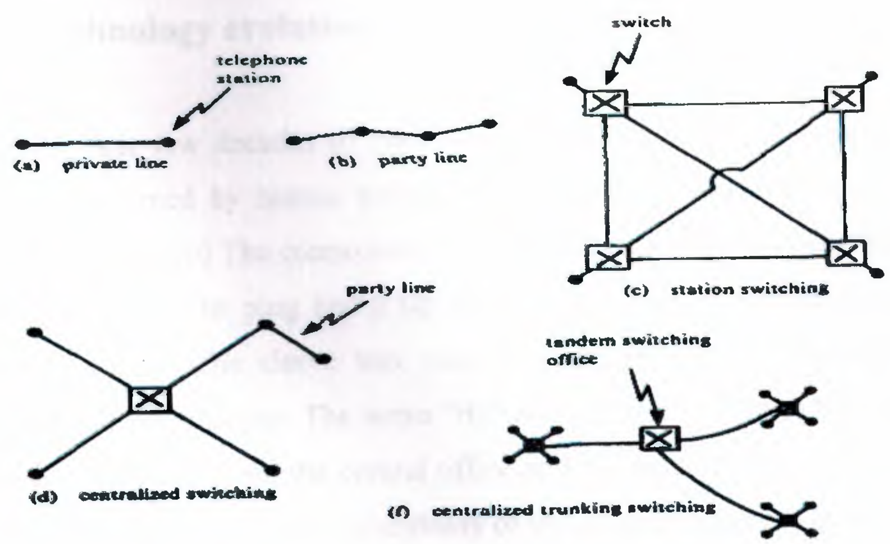
3.10 Switching

3.10.1 Services Evolution

As originally invented by Bell, telephone communication went from a particular telephone instrument to only one other telephone instrument. This was truly private-line service (Figure 3.9 a), and there was no way to reach any other telephones. This private-line service was soon extended to connect a number of telephones to the same line, a form of party line. Everyone could hear everyone else hence there was no privacy, and one call would prevent anyone else from using the line. Clearly, there was the need for a telephone to be switched to any other desired telephone.

One way to perform this switching would be to bring lines from all telephones to all other telephones. A switch at each telephone would then make the connection with the appropriate line to the desired telephone (Figure 3.9 c). If the universe of telephone to be reached is small, such a system would be workable. However, if the universe is large, then the large number of lines that must terminate at each telephone makes such a system of station switching impractical.

The ultimate solution was discovered and implemented only a few years after the invention of the telephone by Bell. The solution was a centralized switching arrangement (Figure 3.9 d). All the lines from all the telephone stations were brought to a common place where the electrical connections were made to connect one station to another. The actual connections were made by people. The central place where the lines all came together was called the "central office" or the "exchange." As exchanges grew to cover greater geographic areas, it became uneconomical to bring lines from the more outlying areas to one central office. More central offices were created, each serving a nearby surrounding area. Connections between central offices were made on lines called *trunks* (Figure 3.9 e). As growth continued, special switching office were developed to handle the trunks between a number of central offices. This centralized switching of trunks was performed at a switching office called a tandem office (Figure 3.9 f)



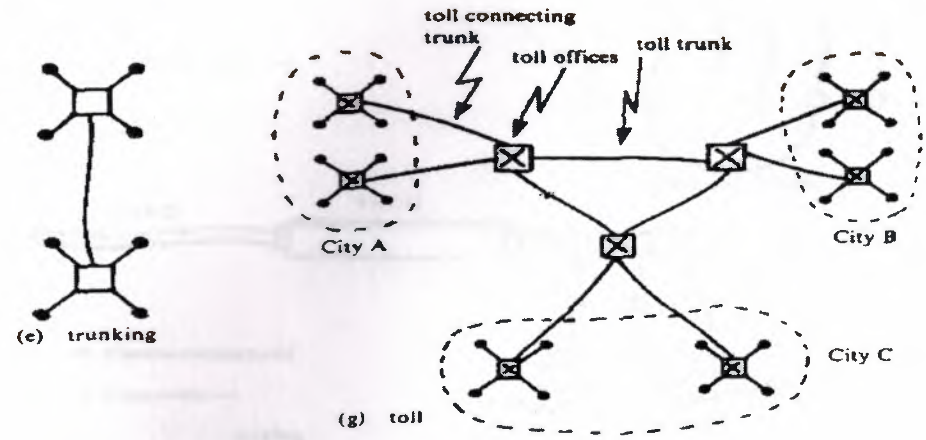


Figure 3.9 Switching Network

New switching needs developed to serve long-distance or toll circuits between cities. Switching offices were thus devised to perform the switching of toll trunks only. These offices were called toll offices (Figure 3.9 g). For businesses, a fair amount of telephone traffic was between telephones that were all located on the customers' premises. These telephones, therefore, could be served most efficiently with a private switch located on the premises. This switch was called a private branch exchange, or a PBX a term that is still in use today. Present PBXs are automatic (PABX), using electromechanical or electronic technology (EPABX).

3.10.2 Technology evolution

During the first few decades of telephone communication, switching was a manual operation performed by human beings who made the actual connections of circuits. (Refer to Figure 3.10) The connections were made at a switchboard utilizing cords with plugs at the ends. The plug had a tip and a ring, which made the actual connection between the lines. The sleeve was used for signaling and supervisory purposes in common-battery exchanges. The terms "tip" and "ring" continue to be used to this day for the two wires between the central office and the actual telephone instrument. The circuits desiring service and the availability of trunks were indicated by small lamps.

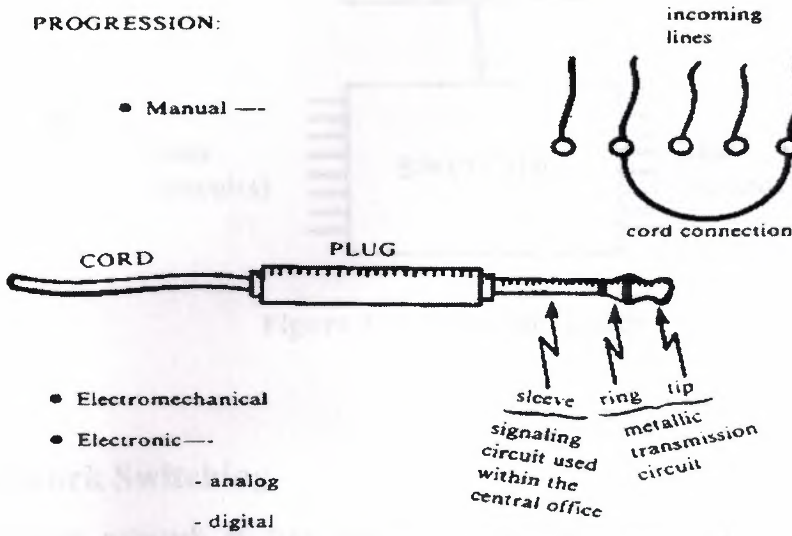


Figure 3.10 Manual Switching System

The first major innovation in switching came in 1892 with the first installation of an automatic switch controlled by the telephone instrument itself. This switch was conceived by Almon B. Strowger. A later modification of this system included the invention of the dial and the use of dial pulses to control the operation of the switching system. The Strowger switch was an electromechanical device. Strowger's invention was adopted for use by the Bell System in 1919. Bell System engineers later developed improved automatic switching systems using electromechanical technology. The electromechanical technology was somewhat slow, not very flexible in terms of offering new services, and frequently generated electrical noise in the connection.

The current generation of technology for telephone switching is electronic using either analog switches or digital switching techniques. The electronic technology is extremely fast and flexible. There are two major parts of any telecommunication switching system: the switches themselves and a control section, as shown by (Figure 3.11) the switches form the switching network. The control section operates the switches at the proper time in order to make a communication connection.

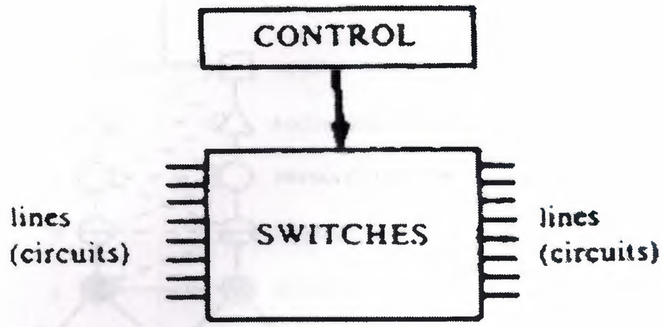


Figure 3.11 Electronic switch

3.11 Network Switching

The switching network is composed of a number of centralized locations called switching offices, where the switching or connecting of circuits is actually performed. These offices are organized into a hierarchy dependent upon whether a call is local (exchange) or long-distance (toll).

The three major types of switching offices arc local, tandem, and toll. The local office is the one closest to the telephone station and connects directly to the local loop. A tandem office serves a cluster of local offices. Toll offices, or toll-centers as they are sometimes called, are concerned with long-distance toll Connections. There are five classes of offices in the Bell System switching hierarchy. At the lowest level is the local office or end office. This is the switching office where the call first originates and finally terminates. The end office is classified as a class-5 office. A class-5 office will attempt to make the connection directly to the terminating class-5 office over a direct interoffice trunk. If necessary, the switching capabilities of a tandem office may be used. If the call is a toll call, then the toll switching offices become involved.

The class-5 offices connect to class-4 offices, or toll centers. Toll centers connect to ciass-3 offices, or primary centers. Primary centers connect to class-2 offices, or sectional centers. Sectional centers connect to class-i offices, or regional centers. The whole hierarchy is organized in a tree fashion with regional centers at the top of the tree, as illustrated in Figure 3.12.

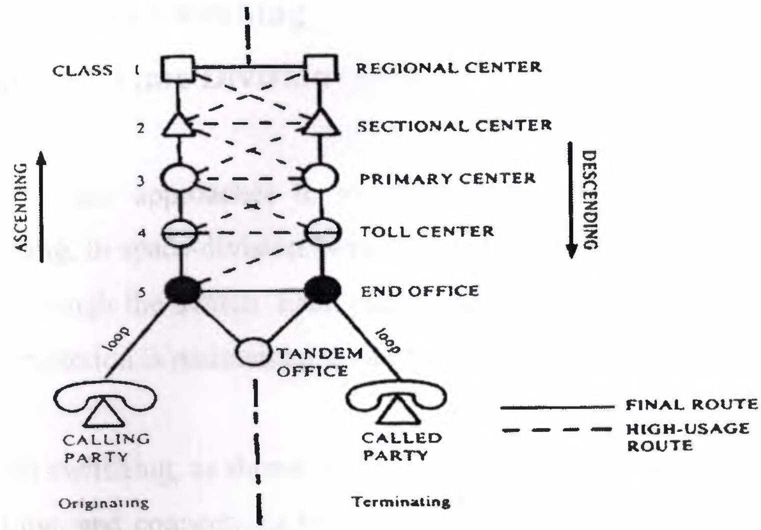


Figure 3.12 Network Switching

The number of offices increases as the tree is descended. There are only about 20 regional centers serving the United States and Canada, while there are well over 20,000 end offices.

Normally, a call will be connected using as few switching offices as possible. The intent is to keep the actual, final route low in the switching network. The preferred route is to connect across from an originating office to a lower terminating office using routes called high-usage routes. If this route is busy, then the next preferred route is to connect across to a center at the same level, or, if this route is busy, then across to the next higher center. If all the routes across are busy, then the route of last resort is to ascend higher in the originating office to the next higher office and attempt to cross over again. The routes connecting centers in the originating offices and in the terminating offices are called intertoll trunks.

The call ascends the hierarchy in the calling-office sequence and then descends the hierarchy in the called-office sequence. Actual traffic conditions at the time of a call determine the final route. The routing used in the AT&T long-distance network is moving toward a dynamic nonhierarchial system in which all switching centers are equal and the route for each call is chosen according to the capacity that is available. Thus, with this newer system, the final routes are not fixed, but are chosen dynamically at the moment of the call. This system offers increased flexibility and efficiency in the use of the network.

3.12 Approaches to Switching

3.12.1 Space and Time Division

There are two basic approaches to switching: space-division switching and time-division switching. In space-division switching, each telephone conversation has its own physical path through the switch. Each path is dedicated to that conversation only, and the physical connection is maintained through out the duration of the conversation.

In time-division switching, as shown by Figure 3.13 paths for separate conversations are separated in time, and connections between various paths are made for very short time intervals. In effect, each conversation is broken into samples, and these sample values are routed to their appropriate destinations along the shared path. This type of switching is most appropriate for signals that have been sampled and digitized.

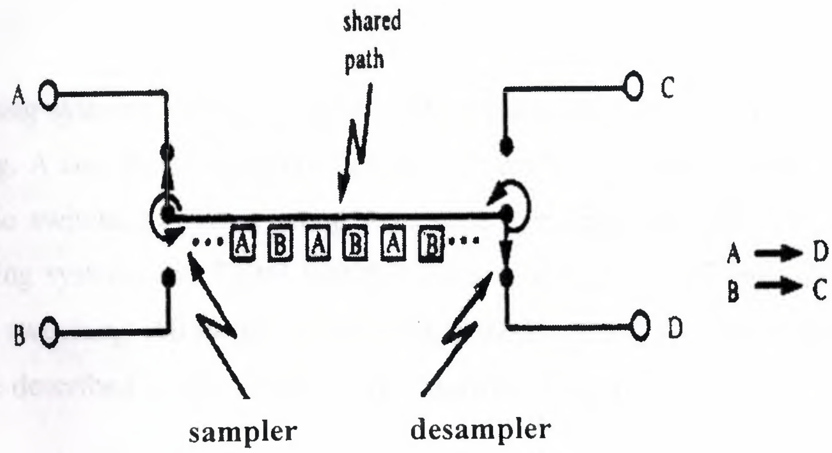


Figure 3.13 Time Division Switching

With space-division switching, signal paths are switched in physical space. With time-division switching, sample values of a number of signals sharing a common medium are reorganized, or switched, in their time sequence. The earliest automatic switches transferred physical paths and were space-division switches. The newest switches are called digital switches because they switch digitized signals by using a combination of space-division and time-division switching.

There is a third possible approach to switching, namely, frequency-division switching. With frequency-division switching, frequency bands would be assigned to different connections. This type of switching could be particularly applicable to the single broadband medium that might be shared by a number of users, such as the coaxial cable in CATV system. Each pair of users would need to agree on a specific band of frequencies to be used for their particular conversation.

3.12.2 Technologies

The two major types of technologies used in switching systems are electromechanical switching and electronic switching. Electromechanical switching is used solely with space-division switching. The earliest type of electromechanical switching system was the step. by-step system using the Strowger rotary switch. An intermediate type of electromechanical switching system is the crossbar system, which utilize, coordinate switching.

Electronic switching systems (ESS) can utilize either space-division switching or time-division switching. A coordinate switching network of small reed switches was used in the first electronic switching system, along with stored program control. The newest electronic switching systems are digital switches using combinations of space-division and time-division switching. All switching is accomplished by using solid-state devices. These systems are described in more detail in the following sections.

3.13 Space Division Switching

Normally, not everyone will want to converse with everyone else at the same time. Hence, it is not necessary to design a switching system so that all lines can be connected simultaneously with all other lines. The incoming lines can therefore be concentrated and distributed through a smaller number of switching paths before being expanded at the last stages of the switching process, as illustrated by the flow chart of Figure 5-8. In this fashion, the switching is accomplished in stages consisting of concentration, distribution, and expansion.

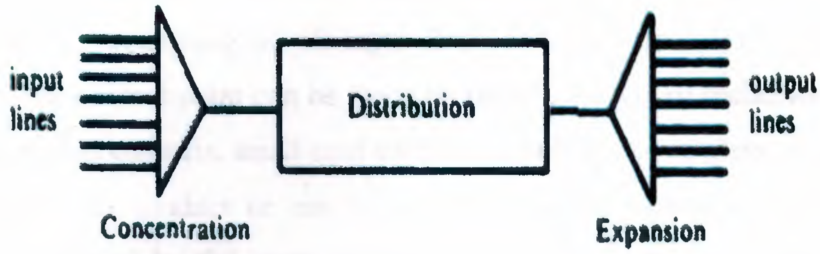


Figure 3.14

Clearly, it is possible that some paths in the switching system might become completely congested. In this case, some calls desiring service will be blocked. The blocking can occur at any stage, and the switching system must be designed to minimize blocking for the traffic that it handles. The actual design depends on the type of customer served by the switching office, insofar as different customers have different traffic patterns.

The two approaches to space-division switching depend on the type of basic switch that is used at each switching stage. The earliest switching systems used the Strowger switch in which a set, or bank, of contacts was swept by one wiper contact. The electromechanical switch invented by Strowger was a stepped rotary switch that could move in two dimensions (see Figure 3.15). The actual electrical connections were made by contacts that wiped across each other and were, therefore, subject to considerable wear and tear. The large amount of mechanical motion meant that a fair amount of time was needed to make the actual, final connection. The Strowger type of switch made a connection between one contact to one of many contacts.

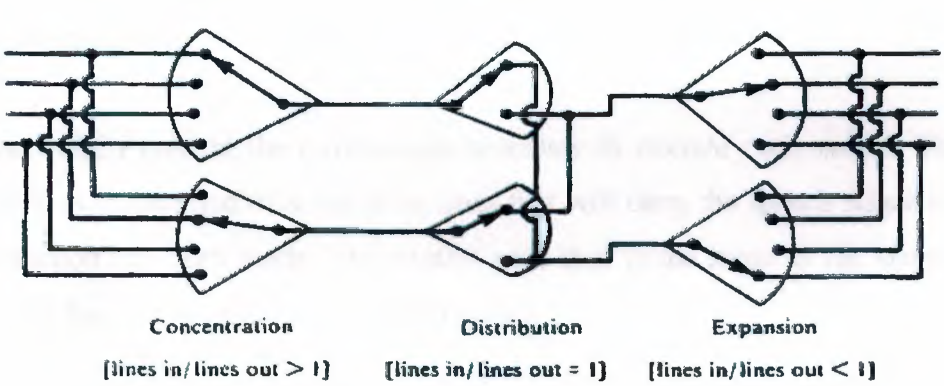


Figure 3.15 Rotary switching network.

An improved approach to switching compared to rotary switching is coordinate, or matrix, switching. In coordinate switching, connections are made at single contact points in a matrix consisting of all input lines and all output lines. The electrical connection at the contact point can be made by using a variety of technologies, such as conventional switch contacts, small reed switch contacts sealed in glass, and diodes and transistors biased to conduct or not to conduct. Many telephone calls are to other telephone lines served by the same switching office. In these cases, the output lines from the switching system are connected to the incoming lines to make the actual connection. Such lines used for intraoffice calls are called intraoffice trunks.

3.14 Control method

A number of methods have been devised over the years to control the connections made by communication switching systems. The method of control is an integral part of the actual switching system, but the various methods of control can also be described separately.

3.14.1 Direct Progressive Control

The oldest method of control is direct progressive control. With this method, sketched in Figure 3.16, each stage of switching responds directly to the digits dialed by the calling telephone. As one stage completes its connection, the next stage responds progressively to the dialed digits. The final connection through the switching system is made gradually as each stage receives its digits and progressively sets up the final path through the system.

The dialed digits contain the information necessary to operate each switch. The digit information is transmitted over the same lines that will carry the speech signal after the last connection has been made. The control path thus is the same as the speech path. Each switch has its own control associated with it.

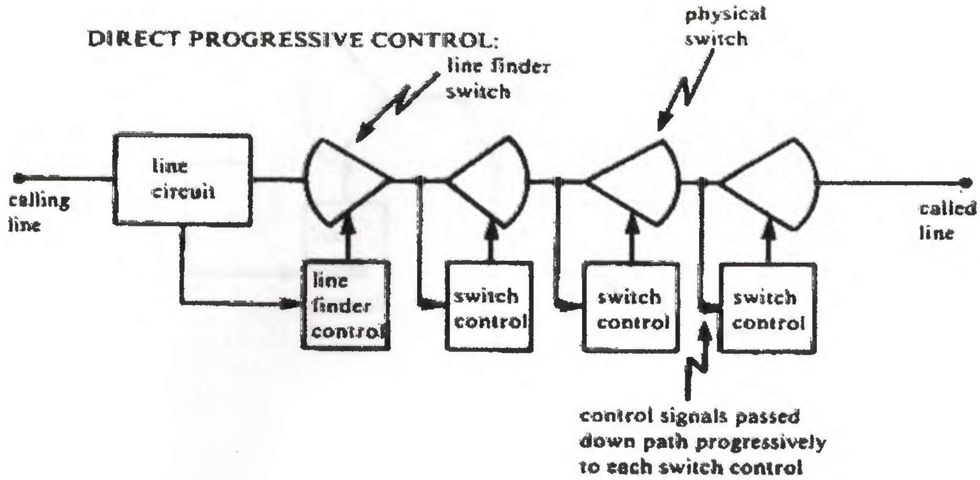


Figure 3.16

Direct progressive control is simple and economical. However, because the paths through the system are created progressively, it is not possible to look ahead to determine whether future paths will be blocked. This means that additional spare capacity is needed to minimize the effects of blocking. Because the dialed digits control the switches directly, there must be a strict correspondence between the dialed digits, which represent the telephone number of the called party, and the physical contact to which the called line is connected.

3.14.2 Register Progressive Control

The solution of some of the disadvantages of direct progressive control is achieved by the use of a register to receive and temporarily store the dialed digits. (Refer to Figure 3.17.) After the called number has been completely dialed into the register, the register then redials the number down the line and into the switches. If blocking occurs, the number is still available in the register so that redialing can be tried. A translator can be used with the register to translate the telephone number into the digits required to control the switches in order to reach the appropriate contact. This means that the actual connection to the switching system does not need to be changed if the telephone number is changed for some reason, and vice versa. The use of a register with progressive control is called register progressive control.

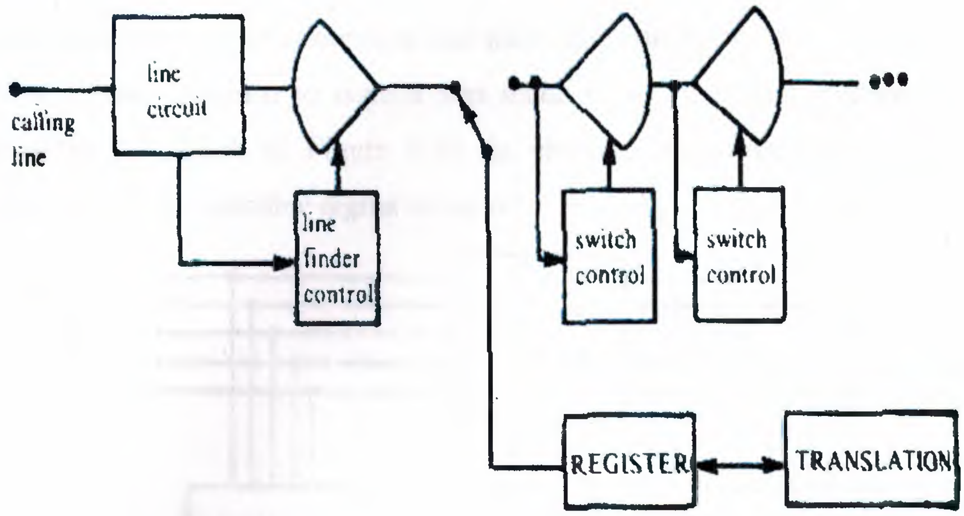


Figure 3.17 Direct progressive controls

Direct progressive control and register progressive control are used in step-by-step switching systems. The basic switch in step-by-step switching systems is the Strowger switch.

3.14.3 Common Control

With progressive control, the path through the switching system is created gradually as each individual switching stage is progressively connected. A more efficient method of control is to determine the optimum path through the switching system and then send signals to each switch over wires that are separated from those used to carry the connected parties. The control mechanism is common to all the switches and is shared by all of them. This type of control, therefore, is called common control. See Figure 5-13. The actual control is performed by circuitry called markers. The markers perform translation, test switches to determine whether they are in use, test different paths, and issue the signals to operate the switches. In effect, the markers are permanently programmed computers based on relay technology. Connectors link registers to the lines desiring service to receive the dialed numbers. The registers are connected to the markers which operate the switches. Common control is used in the crossbar switching systems.

3.14.4 Stored Program Control

With the advent of digital computers and their inherent flexibility by virtue of stored programs, a new approach to control was made possible. Stored program control is illustrated in the sketch of Figure 5-14. In essence, the control of the switches is performed by a programmable digital computer

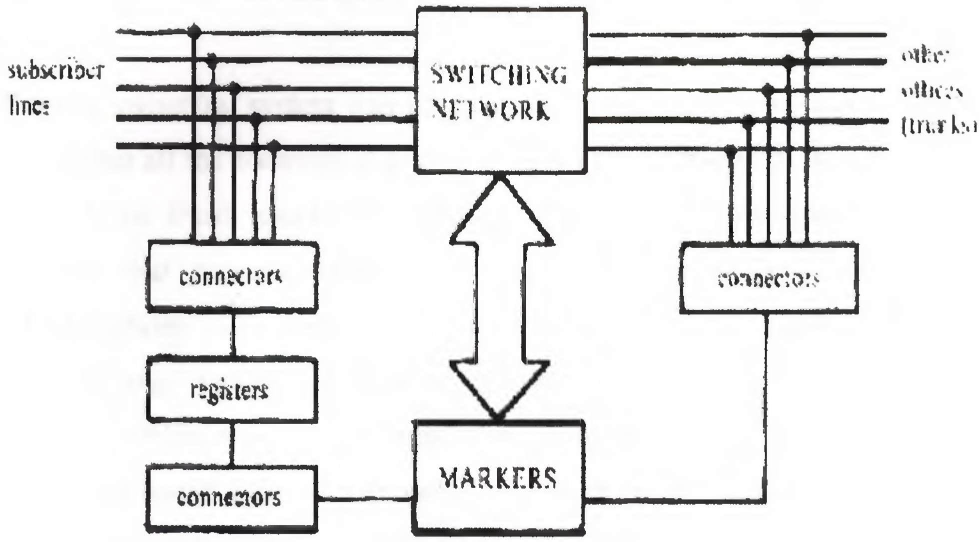


Figure 3.18

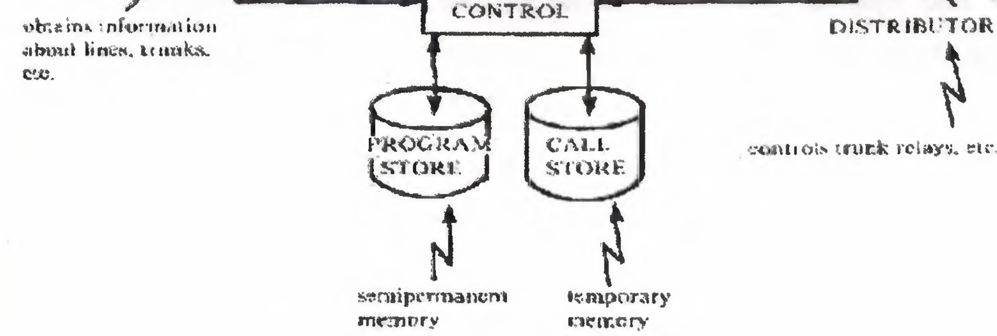


Figure 3.19

information about the dialed numbers. The programs to control the operation of the central control are stored in semipermanent memory called the program store. Jnforitation of a less critical nature is stored temporarily in the call store. Stored program control is used with electronic switching systems.

3.15 STEP-BY-STEP SWITCHING SYSTEM

The step-by-step switching system was invented by Almon B. Strowger, an undertaker who was upset that all the business in his town went to his competitor, whose wife was the operator of the local telephone exchange. The basis of his invention was an automatic switch that stepped vertically and horizontally under the control of digits dialed at the telephone instrument by the telephone user. The first installation of his automatic switch was in 1892. A few years later, in 1896, the telephone dial was invented by some of Strowger's associates. The Strowger switch system was manufactured by the Automatic Electric Company and sold to the non-Bell, independent telephone companies. The first installation of a Strowger switching system in the Bell system did not occur until 1919.

3.15.1 The Strowger Switch

The Strowger switch, shown in the photograph of Figure 5-15, is an ingenious, electromechanical device consisting of electromagnets and ratchets. As we can see from Figure 5-16, the electromagnets are activated by the dial pulses, which cause a wiper assembly to step vertically and then to rotate horizontally in steps until the desired contact is reached. There are 10 vertical steps and 10 horizontal steps, resulting in a total of 100 contacts, any one of which can be accessed. Each horizontal row of contacts is called a bank, with the first bank at the bottom of the switch assembly. The most counterclockwise contact in each bank is called the first contact.

There are two sets of 100-contact banks in a Strowger switch. The bottom set, called the line bank, makes the tip and ring contacts over which the actual speech connection is made. The top set, called the sleeve bank, makes the sleeve contact, which is used for control purposes. A direct current flows through the sleeve contacts to keep the switch operating. When the calling party hangs up, the flow of current through the sleeve circuit ceases, and then springs cause all the switches to return to their rest positions. The action of the wiper across the contacts results in considerable wear on the contacts, as well as a fair amount of noise. There are many mechanical adjustments in the switch, thereby requiring a considerable amount of maintenance. Nevertheless, the basic simplicity and high reliability of the Strowger switch have survived nearly a century of use. Step-by-step switching systems have been supplanted by newer technologies, but step-by-step systems can still be found in use today.

3.15.2 Switching System

All the local-loop wire pairs enter the central office at the cable vault. Cables from the cable vault then terminate at the protector ftame, where a small lightning arrester on each wire pair protects the central-office switching equipment. Wires called jumpers connect from the protector frame to the actual contacts at the input of the switching system.

A step-by-step switching system usually serves 10,000 lines. The switching is divided into stages, illustrated by the flow chart in Figure 5-17. The first stage is the line finder. When the calling party lifts the telephone handset, the telephone instrument starts to draw direct current, which activates relays in the switching system to initiate a hunt by the line finder to locate the line desiring service. The line finder hunts vertically and horizontally until the line desiring service is found. The line finder searches for the 48 volts of common battery that has been placed on the calling line. Each line-finder switch serves 100 lines. Once a particular line finder has been connected to a line, it can no longer serve any other calls. Hence, a number of line-finder switches are needed for each 100 lines, depending on the anticipated traffic.

Each line finder is associated with and wired to the next switch in the sequence of switching stages. This next stage is called the first selector. The selector transmits dial tone to the calling party and then waits for dialing to commence. The first dialed digit causes the selector to move vertically by a number of steps which corresponds to the dialed digit. The selector then searches, or hunts, automatically for an idle line to the next switch.

The next switch in the process is called the second selector. It, too, moves vertically with the dialed pulses and then hunts horizontally for an idle line (a trunk) to the last switch. If an idle trunk cannot be found by either selector, the wiper reaches an eleventh rotary contact, which gives a fast busy tone to the calling party.

The last switch is called the connector. It first moves vertically with the dialed pulses and then rotates horizontally with the last dialed digit. The connector switch performs the final connection to the dialed number.

The above-described switching system can connect to any one of 10,000 lines. Four digits must be dialed by the calling party. The first dialed digit operates the first selector, the second dialed digit operates the second selector, and the last two dialed digits operate the connector. If the switching system were intended to work for seven dialed digits, then three additional selector switches would be needed. If the call were to another office, then the third selector would connect to an interoffice trunk, and the last four dialed digits would operate switches at the distant office.

Chapter 4

SOURCES OF FAILURE IN THE PSTN

What makes a distributed system reliable? A study of failures in the US Public Switched Telephone Network shows that human intervention is one key to this large system's reliability. To operate successfully, most large distributed systems depend on software, hardware, and human operators and maintainers to function correctly. Failure of any one of these elements can disrupt or bring down an entire system. One such distributed system, the US Public Switched Telephone Network (PSTN), is the US portion of possibly the largest distributed system in existence. [1] Like all telephone-switching networks, the PSTN performs a fairly simple task: It connects point A with point B. Paradoxically, this seemingly trivial task requires some of the most complex and sophisticated computing systems in existence. Software for a switch with even a relatively small set of features may comprise several million lines of code.

The PSTN contains thousands of switches. Switches include redundant hardware and extensive self-checking and recovery software. For several decades, AT&T has expected its switches to experience not more than two hours of failure in 40 years [2] a failure rate of 5.7 x 10-6.

Since 1992, telephone companies have been required to notify the US Federal Communications Commission (FCC) of outages affecting more than 30,000 customers. I used these outage records to determine the principal causes of PSTN failures. To account for the possible effects of seasonal fluctuations in call processing volume, I analyzed failures over two years, from April 1992 to March 1994, beginning with the earliest FCC reports. I made quantitative measures of how each failure source affects system dependability, in an effort to shed some light on the dependability of different components (including software).

Major sources of failure were human error (on the part of both telephone company personnel and others), act of nature, and overloads. Overloads caused nearly half of all downtime (44 percent) in terms of outage minutes. An unexpected finding, given the complexity of the PSTN and its heavy reliance on software, was that software errors caused less system downtime (2 percent) than any other source of failure except vandalism. Hardware and software failures were similar in terms of average number of customers affected (96.000 and 118,000) and duration of outage (160 and 119 minutes).

Errors on the part of telephone company personnel and acts of nature caused similar amounts of downtime (14 and 18 percent).

4.1 FAILURE CLASSIFICATION

Table 1 lists the failure classification scheme I used, a scheme that is general enough for comparisons with failures in other large distributed systems. In the case of the human error category, I separated errors made by telephone company personnel from those made by non employees because the companies have direct control over employees only. Overload conditions are accounted for separately because they represent failures accepted as an engineering trade-off between dependability and cost.

Table 4.1 Failure categories

Category	Source	Examples
Human error- -company	Errors made by telephone company personnel	Errors in •cable maintenance • power supply maintenance • power monitoring • facility or hardware hoard maintenance • software version (mismatches) • following software maintenance procedures (such as errors in patch installations an configuration changes; does not include source code changes • data entry
Human error-	Errors made by people	Cable cutting

other	other than telephone company personnel	Accidents (for example, cars striking telephone poles or equipment)
Acts of from nature	Major and minor natural events	Cable, power supply, or facility damaged burrowing animals or lightning
	Natural disasters	Earthquakes, hurricanes, or floods
Hardware failure	Hardware component failure	Failures of cable components, power s or facility components, clock or clock synchronization failure
Software failure	Internal errors in the software	Software errors under normal operation or in. recovery mode
Overload	Service demand exceeds the designed system capacity	
Vandalism	Sabotage or other intentional damage	

4.2 Analysis Procedure

FCC outage reports include the company name and the location of the facility where the outage occurred. They also include the date, time and duration of each outage as well as the number of affected customers. They conclude with a descriptive summary of the outage.

Three of these parameters directly measure failures effects: the number of outages, their duration and the number of customers affected. Using these parameters. I calculated a customer minute's value the number of customers affected multiplied by the outage duration in minutes.

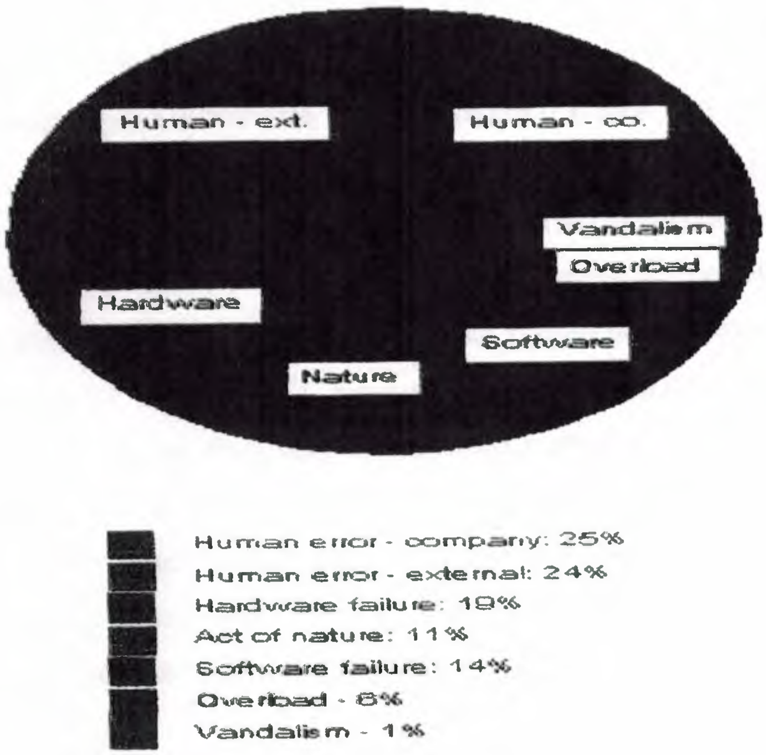


Figure 4.1

Basis for comparing failure data than outage duration alone. For example, a 20-minute outage affecting 10,000 customers (200,000 customer minutes) is considered more severe than a 30-minute outage affecting 1,000 customers (30.000 customer minutes). I did not use an industry measure of outages, user lost erlangs (ULE) [3] because I did not have access to some of the data necessary for computing ULEs. In addition. ULEs are more useful for statistically predicting the duration of future failures, and I wanted to identify and compare the underlying causes of past failures.

I assumed that the FCC reports recorded date and time values at the location where the outage occurred. There is some ambiguity in the times reported: Companies sometimes omitted the time zone, whether it was daylight savings or standard time, and whether the time recorded was an a.m. or p.m. time. The values for customers affected refer to the number of customers served by the failed facility, rather than the customers who were actively using the telephone system at the time of the failure. There may he some variations in the way companies report this value.

I encountered one significant interpretation problem. On April 21, 1992, the Alaskan ocean fiber-cable repeater failed. According to the report, service was restored when the company switched to satellite communications. The company reported the outage duration as two weeks, but it is not clear from the report if it took two weeks to repair the repeater or to switch to satellite communication, although the former appears more probable. L3ec& se this report was unclear, I did not use it in calculation total values or averages. Including overloads as a failure category is somewhat problematic. When an overload occurs, the calls in progress do not fail, but it does prevent the system from accepting additional calls. Since the FCC reports list the number of customers served by an overloaded facility, rather than only the affected customers, these numbers are somewhat misleading. Other types of failure (such as cable cuttings) do prevent service to all customers of an affected facility. Thus, the numbers of customers affected are, in this sense, not directly comparable. The FCC reports do not include the number of customers normally using the system al the time of the outage.

This study excludes overloads when computing failures under he control of the phone companies, because overloads are expected failures.

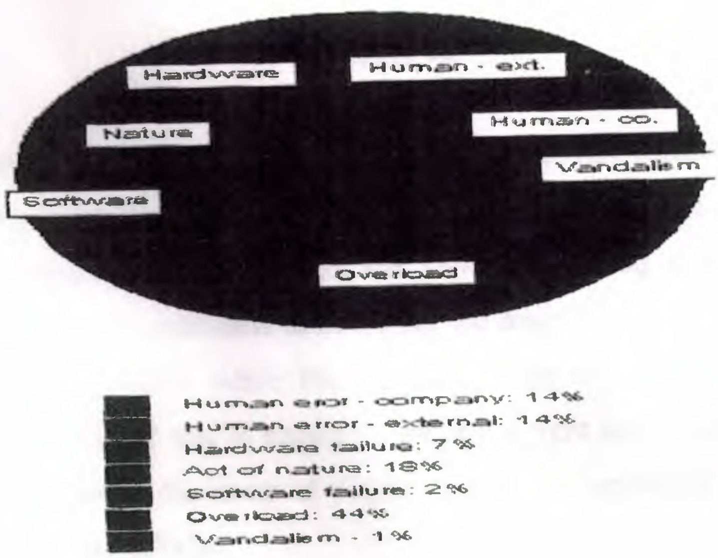


Figure 4.2

4.3 FINDINGS

Table 4.1 summarizes the number and duration of outages, customers affected, and customer minutes by cause. Figure 1 shows the percentage of outages attributed to each major category Figure 2, the percentage of customer minutes. The data show that the number and magnitude of outages differs significantly for most failure categories. For example, although overloads caused only 6 percent of the total outages, they accounted for nearly half the total customer minutes. Human error caused nearly half of the outages, but only about a quarter of the downtime.

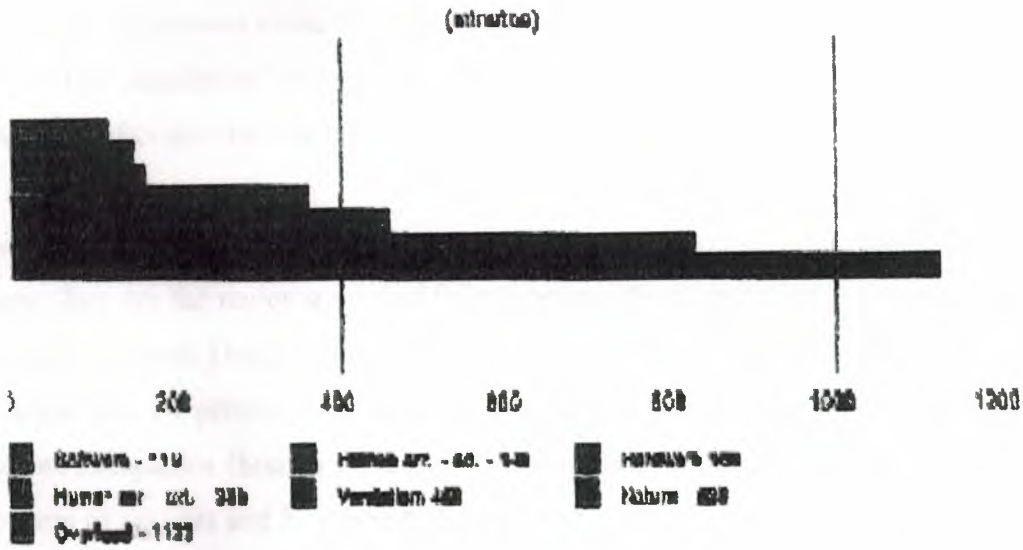


Figure 4.3

Figure 4.3. illustrates the outage durations for the different failure categories and reveals part of the reason number and magnitude measures differ. Software, hardware, and human error company personnel caused the shortest duration outages. Figure 4 compares the duration and customers affected for the major failure categories. The x-axis displays outage duration, while the y-axis displays the number of customers affected. Only overloads and acts of nature (in the upper right corner) are extended and widespread. Failures due to the errors of telephone company personnel (upper left) are brief hut have widespread effects.

Hardware and software failures were similar in terms of outage duration and customers affected. Vandalism and human errors caused by others were also similar in their effects. Table 4.1. Failure effects by categories and sources, for outages from April 1992 to March1994

4.4 OBSERVATIONS

Figure 4.2. shows that nearly half of the downtime is caused by overloads, which are expected outages. Because of economic and technical constraints, telephone companies do not expect service to him available all the time. For example, Bellcore's availability objective for local exchange networks in its client companies is 99.93 percent [4]. Larger capacity networks could probably eliminate most of this downtime but increase cost, through decades of experience, the telephone industry has established a balance between benefits and the cost their consumers find acceptable.

Although the errors attributed to telephone employees are not the major source of outages, they are the major source of failure among those operational aspects under the companies' control. Human error by company personnel accounted for only 25 percent of outages and 14 percent of downtime. However, failure sources controllable by the telephone companies (human error plus hardware and software failures) accounted for 58 percent of outages and 23 percent of downtime.

So human errors by company personnel contributed nearly half of these outages (25 divided by 58) and nearly two-thirds of customer minutes of downtime (14 divided by 23).

Effects of human error were about the same for hardware and software maintenance. Human error for maintenance of cable and hardware components and for power monitoring accounted for about 15 percent of outages and 7 percent of downtime. Software-related human errors included mismatched versions, incorrect data entry, and procedural errors during upgrades. These errors accounted for 10 percent of outages and 7 percent of downtime.

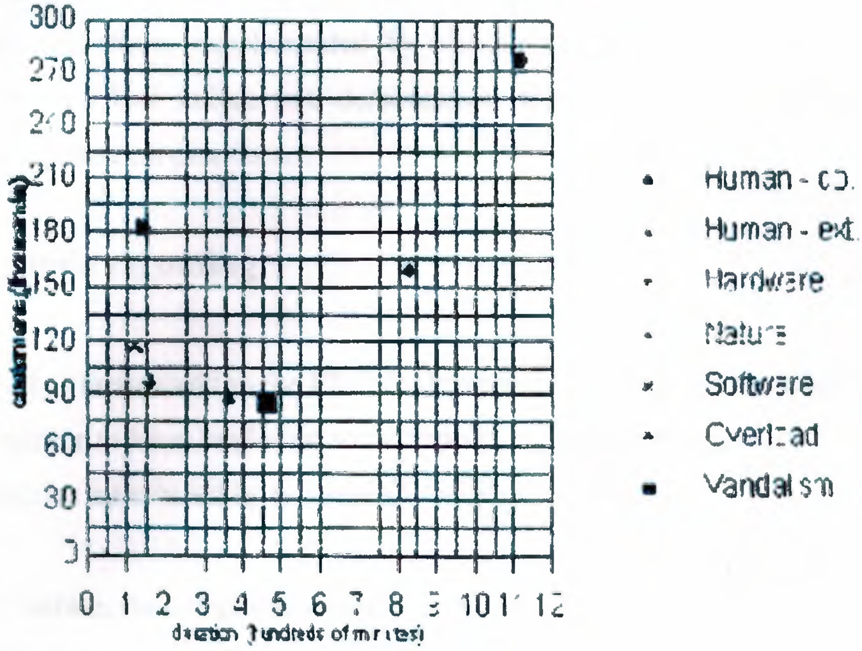


Figure 4.4

Software errors caused a significant number of moderate outages. Although software errors caused approximately 14 percent of the outages, they accounted for only 2 percent of the customer minutes Excluding human error b~ others, acts of nature, and overloads, however, software accounted for 24 percent of outages and 9 percent of customer minutes (downtime). Two factors probably cause software outages to be short: the incorporation of human intervention capabilities in the PSTN and the use of extensive error detection and recovery software.

4.5 WHY SO RELIABLE?

Despite its enormous size and complexity, the PSTN averaged an availability rate better than 99.999 percent in the time period studied. Why should perhaps the world is largest and most complex computerized distributed system is among the most reliable?

4.5.1 Reliable software

To begin with, telephone switch manufacturers are among the world's leaders in computing technology5]. They focus much of their research on developing highly reliable systems. Their software development processes typically incorporate the most sophisticated practices, supplemented by elaborate quality assurance functions. The PSTN software's low failure rate demonstrates that we can develop highly reliable software using the best practices.

4.5.2 Dynamic rerouting

However, other factors add to the PSTN's dependability. In particular, telephone networ designers appear to have exploited some aspects of the network's nature to compensate for complexities introduced by the dependability requirements.

By its very nature, the telephone network is highly distributed, so localized failures are more likely, and switches can reroute traffic dynamically to avoid a failed network node. More important, intermittent failures are usually not catastrophic. Other systems face much greater risks from a failure, no matter how brief For example, failures of a few seconds in some fly-by-wire avionics software may result in the aircraft's destruction. A brief failure in one network component has relatively little impact on the availability figures for t e entire PSTN across the US. However for the PSTN to reroute calls, it must keep a good deal of information global y. Maintaining consistent distributed databases can require complex interactions among system components.

In his book. Normal Accidents, Charles Perrow identified two factors-interactions and coupling- that are significant in determining a system 5 safety properties [6]. Interactions refer to the dependencies between components, while coupling refers to the flexibility in a system. He characterized interactions as linear or complex, while coupling is loose or tight Systems with simple, linear interactions have components that affect only other components that are functionally downstream. Complex system components interact with many other components in different parts of the system. Loosely coupled systems have more flexibility in time constraints, operation sequencing, and assumptions about the environment than do tightly coupled systems. Systems with complex interactions and tight coupling are likely to promote accidents. Complex interactions allow for more complications to develop and make the system hard to understand and predict. Tight coupling also means that the system has less flexibility in recovering when things go wrong.

John Rush by applied Perrows analysis of failures in large physical systems to computer system [7]. In such systems, interactions can, for example, take the form of signaling that coordinates processes or keeps distributed databases consistent. Coupling refers to constraints on timing, operation sequencing, acceptable input data ranges, and other aspects of system flexibility. Control systems with non-negotiable, real-time deadlines are tightly coupled, while the Internet, with multiple paths to route packets, is a loose-coupling example Systems that require frequent updating of a distributed database are likely to have complex interactions to exchange messages among components and maintain the databases global consistency. A simple update and reporting system, which updates a database and writes files for input to report programs, is an example of linear interaction.

4.6 LOOSE COUPLING

In most system, a trade-off can be made between simplicity of interactions and looseness of coupling. We can consider the PSTN a loosely coupled system because it can dynamically reroute calls along many paths. However, it achieves this loose coupling at the cost of some complex interactions between components. These include the need for end-to-end acknowledgments, interactions among many systems, and the maintenance of some globally consistent databases. Major switching centers store information on alternative paths and exchange data on traffic patterns and switch status throughout the day. Such complex interactions can contribute to failures by making system behavior difficult to analyze.

The most spectacular example of a failure due to complex interactions in the PSTN is the 1990 nationwide AT&T network failure. This failure resulted from interactions between systems attempting to maintain consistent information about a failed switch. On the other hand, the PSTN's distributed database of routing information promotes loose coupling, which contributes to system dependability.

For a communications system, coupling is probably the more important of the two properties in determining its capacity to tolerate failures. It is directly related to the systems primary function: maintaining connections between points. The PSTN is loosely coupled, allowing for flexibility in recovering from failures. For the PSTN, loose coupling probably more than makes up for the interaction complexity. Designers should consider the trade-off between these factors-linear interactions or loose coupling-to adds dependability to any high-integrity system. Two levels of recovery mechanisms-automated and manual-exploit the PSTNs loose coupling.

Designers devote about half of the software in telephone switches to error detection and correction. Such a high percentage of self-checking is probably atypical for software systems. Although some researchers note that adding fault-tolerance and fault-avoidance mechanisms to software sometimes decreases dependability because of the recover mechanisms' added complexity [8], these mechanisms work with great success in switching systems. Other computer-driven systems might benefit from more extensive use of built-in diagnostic and recovery software.

4.7 HUMAN INTERVENTION

In addition to built-in self-test and recovery mechanisms, operators monitor telephone switches 24 hours a day and usually have the ability to modify switch software on the fly. Switch manufacturers provide 24-hour Support services, usually with a remote maintenance capability that allows them to correct software in a switch thousands of miles away. Human intervention corrected many failures in under one hour. Simply restarting a switch temporarily fixed a significant number of software-ca used out-ages.

Traffic routing also benefits from automated and human operations. Using information on switch status and traffic patterns exchanged by switches, software within a switch will automatically select an alternative route if the preferred route becomes overloaded or unavailable If the switch exhausts all alternative routes, human intervention can reconfigure the network, sometimes solving the problem in a few minutes. Status data exchanged regularly between switches makes automated and human operations to reconfigure routing possible. PSTN designers made the coupling-interactions trade-off in favor of loose coupling. Loose coupling allows human operators to intervene in the event of failure, rather than relying entirely on computer control.

Software is not the weak link in the PSTN system's dependability. Extensive use of built-in self-test and recovery mechanisms in major system components (switches) contributed to software dependability and are significant design features in the PSTN. The network's high dependability indicates that the trade-off between dependability gains and complexity introduced by built-in self-test and recovery mechanisms can be positive. Likewise, the tradeoff between complex interactions and loose coupling of system components has been positive, permitting quick human intervention in most system failures and resulting in an extremely reliable system.

CONCLUSION

Despite its enormous size and complexity, the PSTN averaged an availability rate better than 99.999 percent in the time period studied. Why should perhaps the world's largest and most complex computerized distributed system also be among the most reliable. To begin with, telephone switch manufacturers are among the world's leaders in computing technology [5]. They focus much of their research on developing highly reliable systems. Their software development processes typically incorporate the most sophisticated practices, supplemented by elaborate quality assurance functions. The PSTN software's low failure rate demonstrates that we can develop highly reliable software using the best practices.

Since the 1970, communication technology has made great advancement offering capabilities, which were impossible only a decade earlier. The bandwidths of new transmission technologies are virtually unlimited, offering the ability to carry numerous channels of voice, data, and video signals. The capacity and speed of the new switching technologies are staggering, while their actual physical size has shrunk. Simultaneous with these technologies advance, the whole communication industry has undergone dramatic restructuring.

A circuit switched network, such as the PSTN. provide end to end connection on demand, as long as the necessary network resources are available the connections end to end delay is usually small and always constant, and other user's cannot interfere with the quality of communication.

REFERENCES

1. Theodore's S. Rappaport. (1993). Telecommunication. Prentice Hall.

2. Mertin Nemzow. (1999). Implementing PSTN McGraw-Hill.

3. Vigaay Gag, Joseph Vilkes. (1996). PSTNand Personal communication systems.

4. Joseph Schiller. (2000). Mobile communication. Prentice Hall.

5. PSTN Communication. Raymeed. (1990).