BACHELOR OF COMPUTER APPLICATIONS

SEMESTER *5*

DCA3101

WEB DESIGN

Unit 3

Internet Services & Internet Security

Table of Contents

INSPIRED BY LIFE

1. INTRODUCTION

In previous unit, you have learned the concepts of World Wide Web. The purpose of WWW is to display the data in an organized way.. In this unit you are going to study protocols, Internet services like email, FTP, newsgroups, and Internet security issues and measures to be taken. Finally, you will also studydomain name system and IP addressing concepts.

1.1 Objectives:

After studying this unit, you should be able to:

- *Describe internet services*
- *Explore Networking protocols*
- *Comprehend internet security issues*
- *Describe domain name message format*
- *Explain IP addressing*

1.2 Networking Protocols

The internet uses different standards to transfer information to and from remote computer networks. These standards known as protocols, allow computers to communicate with one another in a structured method.

The Widely used internet protocols are:

- **Transmission Control Protocol**

  Transmission Control Protocol (TCP) is a well-known communication protocol used for network communication. Any message is divided into a number of packets that are transferred from the source to the destination, where they are then reassembled.

- **User Datagram Protocol (UDP)**

  UDP is a communication protocol used largely for generating loss-tolerant and low-latency connection between various applications. It is an alternative for Transmission Control Protocol.

- **HTTP**

  The Hypertext Transfer Protocol (HTTP) is a networking protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web.

- **Hyper Text Transfer Protocol Secure (HTTPS)**

  HTTPS is a standard protocol that protects communication between two computers, one of which is using a browser and the other is requesting data from a web server. the data transfer taking place in an encrypted fashion it prevents hackers from deciphering or altering data while packets are being transferred.

- **FTP (File Transfer Protocol)**

  FTP is probably the most used resource on the Internet. It is a protocol which allows users on computer to transfer files to another computer or we can say to exchange files over the Internet. FTP works in the same way as HTTP (Hypertext Transfer Protocol) for transferring Web pages from a server to a user's browser and SMTP for transferring electronic mail across the Internet in that, like these technologies, FTP uses the Internet's TCP/IP protocols to enable data transfer.

  When you take a file from a remote machine down to your own machine, it is called a download. When the opposite happens and you place a file on a remote machine, it is called an upload. FTP is most commonly used to download a file from a server using the Internet or to upload a file to a server.

- **SMTP (Simple Mail Transfer Protocol)**

  SMTP is used for sending E-mail messages between servers. Most e-mail systems that send mail over the internet use SMTP to send messages from one server to another, the messages can then be retrieved with an e-mail client. In addition, SMTP is generally used to send messages from a mail client to mail server.

- **Telnet**

  A series of protocols known as Telnet are used to link up systems. Here, the method of connecting is known as remote login. The local computer is the system that makes the connection request, while the distant computer is the system that grants the connection.

- **Gopher:** Gopher is a set of rules used for retrieving, displaying, and searching content from remote locations. Client/server technology is used in Gopher as well.

1.3 Internet Services

Internet Services refers to the exchange of data communications between computers using a standard network protocol over the public and private, interconnected network.

Some of the internet services are:

- **Electronic Mail**

  Electronic mail could be defined as the transmission of letters and memos from one computer to another. When E-Mail originated in the 1970s, it was used for sending textual messages only. The capability to send various items has rapidly changed as now E-Mail users can attach various documents like spreadsheets, business forms, lengthy documents, scanned images, faxed images, computer graphics, meeting schedules, sound, and video to their messages (and the list continues to expand).

E-Mail Services

In practice, E-Mail usually refers to a service that includes the following facilities:

- **Store-and-forward:** Messages are held until they are requested by the recipient. Direct person-to-person contact is not required, and the service can be used by either party at whatever time and on whatever day that suits them.
- **Blind copies:** Copies can be sent automatically to the names on a distribution list, including 'blind' copies (where the principal recipient is not notified that others have received the message).

- **Advise delivery:** The sender can be told (by a confirming message to his or her mailbox) when the recipient has read the message. An immediate reply could also be demanded.
- **Off-line working:** Text can be prepared in advance of transmission, and incoming messages can be saved for later consideration or for use within word-processed documents.

➢ **File Transfer Protocol (FTP)**

Transferring data from one host to another is one of the most frequently used operations. Both needs to upload data (transfer data from a client to a server) and download data (retrieve data from a server to a client) which is addressed by FTP. Additionally, FTP provides security and authentication measures to prevent unauthorized access to data.

Application of FTP:

1. Every webmaster uses FTP to transfer web pages, web application files and images to their web server.
2. Corporations uses FTP server to provide common files to their clients and branch offices.
3. One huge advantage to using FTP is that there is no restriction on the file size.

➢ **Newsgroups**

Newsgroups started out as electronic bulletin board where people posted questions and answers. Most local communities have bulletin board in public places; for example, grocery stores, student union buildings often have bulletin boards on which individuals can post notices about events, items for sale, properties for rent, want ads, and posters about causes or issues they support. These bulletin boards provide a venue where people can drop in to post information or read what others have posted. You do not have to belong to a group to read or post bulletin board message. You just need access to the bulletin board. News group are online bulletin boards first made available in the 1980s over the distributed user network.

- **Other Internet Services**

Discussion Groups

Discussion groups are similar to Newsgroups in the manner in which they work. However, there is one significant difference. In a discussion group, the participants are normally restricted to a private entity like a corporation, a training or customer support group, or a special interest group or class. The newsgroup is open to internet while a discussion group may have password protections, be part of a corporate internet, or part of some virtual private network to limit its access to those who are registered to the discussion group.

Internet chat

Chat rooms work with a variety of mechanisms. The most common is the Internet Relay Chat (IRC) protocol which is one of the TCP/IP protocols. The chat defines a channel (Chat room) that will send all the messages typed by one member of the channel to all the other clients logged into that channel (chat room). Some chats require special browsers and charge fees on monthly basis. Most are moderated and have posted rules of conduct that allows them to deny service to anyone who does not conform to the rules.

SELF-ASSESSMENT QUESTIONS - 1

1. _____ is defined as the transmission of letters and memos from one computer to another.
2. FTP stands for_____ .
3. Which of the following internet service is similar to newsgroup

   a) Email b) video sharing c) discuss group d) online gaming

2. THE INTERNET SECURITY

The internet is a vital resource that is changing the many enterprises and individuals communicate and do business. However, the internet suffers from significant and widespread security problems. Many agencies and enterprises have been attacked or probed by intruders, with resultant losses to productivity and reputation. In some cases, enterprises had to be disconnected from the Internet temporarily, and have invested significant resources for correcting problems with system and network configurations. Sites that are unaware of or ignorant of these problems face a risk that network intruders will attack them. Even sites that do observe good security practices face problems with new vulnerabilities in networking software and the persistence of some intruders.

The fundamental problem is that the Internet was not designed to be very secure. Some of the Internet problems are:

Ease of Eavesdropping and spoofing

The majority of the Internet traffic is not encrypted. Email, passwords, and file transfers can be monitored and captured using readily available software.

Vulnerable TCP/IP Services

A number of the TCP/IP services are not designed to be secure and can be compromised by knowledgeable intruders. Services used for testing are particularly Vulnerable.

Lack of Policy

Many sites are configured unintentionally for wide open Internet access without regard for the potential of abuse from the Internet. Many sites permit more TCP/IP services than they require for their operations and do not attempt to limit access to information about their computers that could prove valuable to intruders.

Complexity of configuration

Host security access controls are often complex to configure and monitor. Controls that are accidentally misconfigured can result in unauthorized access.

Malware

It is often known as "malicious software," is a term that refers to a variety of harmful software, such as Trojan horses, worms, and computer viruses.

Computer worm

A computer worm is a piece of software that replicates itself on other computers. These copies can spread quickly and in large quantities without the need for human interaction.

Spam

Unwanted emails in your inbox are referred to as spam. Junk mail that promotes products or services you don't want to buy can occasionally be considered spam. Though most of these are regarded as being safe, some of them may contain links that, if opened, will infect your machine with harmful malware.

Phishing

Cybercriminals use phishing scams to obtain confidential or sensitive information. They might pretend to be your bank or an online service, enticing you to click links that ask you to confirm passwords or account information.

Botnet

A botnet is a collection of infected private computers. These machines are possessed by a single person, are infected with malicious software, and are frequently persuaded to carry out evil deeds like spamming or denial-of-service (DoS) attacks.

2.1 E-Commerce Security Issues

Today's world is digital information society, and high-tech technologies especially electronic information technology has given great impact to all walks of life. E-commerce was generated in such an information age and is developing rapidly; it has gradually become a new model for business activities.

E-commerce security is the protection of e-commerce assets from unauthorized access, use, alteration, or destruction.

Security tools for e-commerce

- Firewalls
- Digital certificates
- Public key infrastructure
- Encryption software

Firewall

A firewall is a software program or piece of hardware that helps screen out hackers, viruses, and worm that try to reach your computer over internet.

Digital certificate

It is an attachment to an electronic message used for security purposes. The most common use of a digital certificate is to verify that a user sending a message is who he or she claims to be, and to provide the receiver with the means to encode a reply.

Public key infrastructure (PKI)

PKI is a security architecture that has been introduced to provide an increased level of confidence for exchanging information over an increasingly insecure Internet.

Encryption

Encryption is a process which is applied to text messages or other important data, and alters it to make it humanly unreadable except by someone who knows how to decrypt it.

Encryption is the most effective way to achieve data security. To read an encrypted file, you must have access to a secret key or password that enables you to decrypt.

2.2 The Internet Security Measures

Data protection on the internet demands the use of a variety of tools and technology. While taking the right steps to help keep your network secure, it's necessary to take a variety of internet security tactics into consideration. These strategies may consist of:

- **Browser selection:** Although each browser has its own security mechanisms in place, some of them may have severe weaknesses that hackers and other cybercriminals can take advantage of and penetrate. To decrease the possibility of your computer or network being compromised, make sure you're using a secure browser.
- **Multi-factor authentication (MFA):** MFA is a technique for restricting access to computers by demanding several independent pieces of proof for an authentication process. Having a user provide at least two authentication factors can increase the security of websites and email accounts.
- **Email security:** Viruses, worms, Trojan horses, and other undesirable programs benefit greatly from email traffic. A multi-layered, all-encompassing email security plan will greatly limit susceptibility to new threats. Moreover, emails can be secured using cryptography by being signed, having their bodies encrypted, and having their communications between mails servers encrypted.
- **Firewalls:** Firewalls serve as filters that safeguard equipment by approving or rejecting network access. Firewalls can stop the theft of confidential data and the embedding of malicious code onto networks by using a specific set of rules to determine whether something is safe or destructive.

INSPIRED BY LIFE

SELF-ASSESSMENT QUESTIONS - 2

4. Which of the following is not internet problem with TCP/IP
   i) Ease of eavesdropping and spoofing
   ii) Vulnerable TCP/IP services
   iii) Firewall
   iv) Complexity of configuration
5. To read an encrypted file, you must have access to a _____or_____ that enables you to decrypt.

INSPIRED BY LIFE

3. DOMAIN NAME SYSTEM (DNS)

On a TCP/IP network, computers know each other by their IP addresses. But for human beings, remembering numbers is not the easiest thing to do. Remembering names is much easier. Similarly, a way was devised to associate IP addresses with names that can be easily remembered. In the early days of the Internet, "hosts" files were used to associate machines with names. The hosts file is simply a table of IP addresses and corresponding names like a phone directory. Any name lookup (the process of identifying the IP address associated with a name) will first check the hosts file (if present) on the machine making the query, to see whether the name can be resolved.

Within the Internet, each separate computer is called a host. For example, you might tell someone he can find the information he wants by connecting to a host in Switzerland. If your computer is connected to the Internet, then it too is a host, even though you may not be sharing any resources with the rest of the world. If you connect to and log into a host and then use its functions to reach out onto the Internet, you are using your computer as a terminal to reach another computer. Host connections are designed to use very simple text based interactions.

Being connected to the Internet means your computer system or network is actual a node on the Internet. It has an individually assigned Internet address and client program to in running on the computer system that can take full advantage of the computer's capabilities. Your workstation is a peer of every other computer on the Net. So, a node is any "addressable device" attached to a computer network.

But with the number of hosts on the Internet increasing rapidly to an unmanageable level, that soon became impossible. The way out was the DNS: the Domain Name Server. The DNS is a distributed, scalable database of IP addresses and their associated names. It is distributed in the sense that unlike the hosts file, no single computer contains all the DNS information in the world. The DNS data is distributed across many name servers. It is scalable – you can increase the volume of total DNS data and requests from machines for the same data, without significantly increasing the querying time. Otherwise the World Wide Web would really become the World Wide Wait.

To understand the DNS and the way it is used, we need to understand the Internet naming structure.

for example, the address: http://www.trg.hclsso.hclinfosystems.com/

www: Indicates that the machine is part of the world

com: Indicates the top-level domain (TLD) that the machine is part of. Top Level Domain include .com, .edu, .gov, .in etc

hclinfosystems: Shows that the computer we are looking for is in a network called hclinfosystems

hclsso: Indicates a sub-network (a group of computers with a common function or at a common location).

trg: Is the name of the machine that we are interested in.

Let us see how the DNS aids in identifying the machine's IP address, given its name at the top level of DNS structure are the nine root name servers of the world, which contain pointers to the master name servers of each of the top-level domains. To find the IP address of http://www.trg.hclsso.hclinfo systems.com/ the DNS server will have to ask one of the root name servers for the address of the master name server for the .com domain. This master name server will have the addresses of the name servers for all the .com domains. From here you get the address of the name server, for the hclinfosystems.com/ domain. You move on to this name server, which will give you the IP address of the machine trg.hclsso.com. If there is a name server for the trg.hclsso.com sub-domain, then the name server for hclinfosystems will guide you on to this name server, which will give you the IP address of trg.

A domain name is a way by which a company can uniquely identify itself on the Internet. Registering a domain name on the Internet is the equivalent of registering a company name at Companies House. Based on the top level identifications, there are basically two types of domains:

1. Non-geographic domains
2. Geographic domains

Non Geographic Domains

The top level Internet domain types those are non-geographical:

Domain	*Indicates*	*Example*
Com	Commercial Organizations	hclinfosystems.com
Edu	Educational Institutions	Stanford.edu
Mil	A (US) military setup	Nic.mil
Gov	A (US) government setup	Nasa.gov
Org	Other organizations	www.bjp.org
Net	Other networks	Ns.stph.net
Int	An international organization	Tpc.int

Geographic Domains

The geographically based top-level domains use two-letter country designations.

Domain	Meaning
Au	Australia
Ca	Canada
Dk	Denmark
Fr	France
Gr	Greece
In	India
Jp	Japan
Us	United States

In a complete (fully qualified) domain name, the part furthest to the right is the top level domain, representing either a type of organization or a country. As you read in from the right, the name gets more specific until you reach the name of the individual host computer. For instance: rubens.anu.edu.au is the name of a computer. It is in Australia (au), in the educational area (edu), at the Australian National University (ANU) and the host computer is named rubens.

3.1 IP Addressing

IP addresses are represented by a 32-bit unsigned binary value. It is usually expressed in a dotted decimal format. For example, 9.167.5.8 is a valid IP address. The numeric form is used by IP software.

The IP Address

To identify a host on the Internet, each host is assigned an address, the IP address, or in some cases, the Internet address. When the host is attached to more than one network, it is called multihomed and has one IP address for each network interface. The IP address consists of a pair of numbers:

IP address = <network number><host number>

The network number portion of the IP address is administered by one of three Regional Internet Registries (RIR):

1. **American Registry for Internet Numbers (ARIN):** This registry is responsible for the administration and registration of Internet Protocol (IP) numbers for North America, South America, the Caribbean, and sub-Saharan Africa.
2. **Reseaux IP Europeans (RIPE):** This registry is responsible for the administration and registration of Internet Protocol (IP) numbers for Europe, Middle East, and parts of Africa.
3. **Asia Pacific Network Information Centre (APNIC):** This registry is responsible for the administration and registration of Internet Protocol (IP) numbers within the Asia Pacific region.

IP addresses are 32-bit numbers represented in a *dotted decimal* form (as the decimal representation of four 8-bit values concatenated with dots). For example, 128.2.7.9 is an IP address with 128.2 being the network number and 7.9 being the host number. Next, we explain the rules used to divide an IP address into its network and host parts.

The binary format of the IP address 128.2.7.9 is:

10000000 00000010 00000111 00001001

IP addresses are used by the IP protocol to uniquely identify a host on the Internet (or more generally, any internet). Strictly speaking, an IP address identifies an interface that is capable of sending and receiving IP datagrams. One system can have multiple such interfaces. However, both hosts and routers must have at least one IP address, so this simplified definition is acceptable. IP datagrams (the basic data packets exchanged between hosts) are transmitted by a physical network attached to the host. Each IP datagram contains a source *IP address and a destination IP address*. To send a datagram to a certain IP destination, the target IP address must be translated or mapped to a physical address. This might require transmissions in the network to obtain the destination's physical network address.

Class – Based IP Addresses

The first bits of the IP address specify how the rest of the address should be separated into its network and host part. The terms *network address* and *netID* are sometimes used instead of network number, but the formal term, used in RFC 1166, is network number. Similarly, the terms host address and *hostID* are sometimes used instead of host number.

There are five classes of IP addresses. They are shown in Fig.3.1.

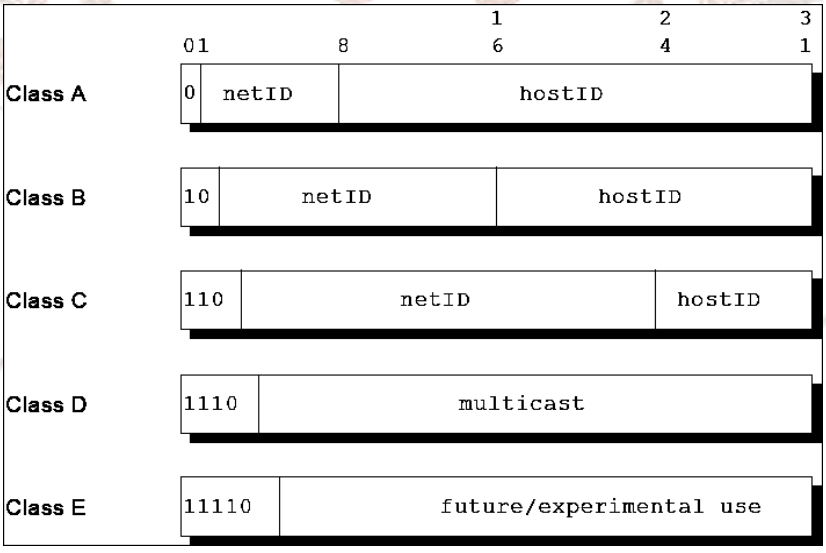

Figure 3.1: IP: Assigned classes of IP addresses

Where:

Class A addresses: These addresses use 7 bits for the <network> and 24 bits for the <host> portion of the IP address.These are very big networks with up to 224 (16 million) nodes. Class A networks have their network addresses from 1.0.0.0 to 126.0.0.0. The zeros are replaced with the node addresses. NEARNET, Sprint, ANSnet, Merit and AT&T are examples of organizations with class A network numbers.

Class B addresses: These addresses use 14 bits for the <network> and 16 bits for the <host> portion of the IP address. Class B networks are smaller than Class A networks. They can have up to a maximum of 65000 nodes. Network addresses range from 128.0.0.0 to 191.0.0.0. In this case only the last two zeros are replaced with the node addresses. Class B addresses go to organizations with larger nets, such as universities or large businesses.

Class C addresses: These addresses use 21 bits for the <network> and 8 bits for the <host> portion of the IP address. Class C networks are smaller than Class B networks. They can have up to 254 nodes. Network addresses range from 192.0.0.0 to 223.0.0.0. In this case only the last zero are replaced with the node addresses. This class is where most networks will be assigned.Originally,Class C addresses were intended for small company networks, K-12 schools and single machines that were not connected to other, larger nets.

Class D addresses: These addresses are reserved for multicasting (a sort of broadcasting, but in a limited area, and only to hosts using the same Class D address).

Class E addresses: These addresses are reserved for future or experimental use.

Class A address is suitable for networks with an extremely large number of hosts. Class C addresses are suitable for networks with a small number of hosts. This means that medium-sized networks (those with more than 254 hosts or where there is an expectation of more than 254 hosts) must use Class B addresses. However, the number of small- to medium-sized networks has been growing very rapidly.

3.2 Reserved IP addresses

A component of an IP address with a value all bits 0 or all bits 1 has a special meaning:

6. All bits 0: An address with all bits zero in the host number portion is interpreted as *this* host (IP address with <host address>=0). All bits zero in the network number portion is *this* network (IP address with <network address>=0). When a host wants to communicate over a network, but does not yet know the network IP address, it can send packets with
   <network address>=0. Other hosts in the network interpret the address as meaning this network. Their replies contain the fully qualified network address, which the sender records for future use.
7. **All bits 1:** An address with all bits one is interpreted as *all* networks or *all hosts*. For example, the following means all hosts on network 128.2 (Class B address): 128.2.255.255.
8. **Loopback:** The Class A network 127.0.0.0 is defined as the loopback network. Addresses from that network are assigned to interfaces that process data within the local system. These loopback interfaces do not access a physical network.

Subnet Masks

The subnet mask is used by the internet layer to determine which part of the IP address is the network ID and which part is the host ID. The subnet mask also can be used to determine whether a subnet is defined and to find the ID of that subnet. The TCP/IP subnet mask specifies that the octets of the IP address marked as 255 are the network ID and octets marked by 0 are the host ID. Any part of the subnet mask with 1s specifies the network portion of the address. 0s in the subnet mask specify the host portion of the address. The 1s are always at the first of the subnet mask, because an IP address always specifies the network portion of the address first. The host ID is specified by the remaining numbers of the IP address, which correspond to the 0s at the end of the subnet mask. In a subnet mask, note that the 1s are always grouped together and the 0s are always grouped together. The subnet mask basically divides the IP address into two pieces: the network ID and the host ID.

The computation TCP/IP performs is a logical bitwise "AND" of the IP address and the subnet mask. The calculation sounds complicated, but all it really means is that the address in its true 32-bit binary format is logically "ANDed" with the subnet mask (also a 32-bit binary number). This extracts the network ID.

The default subnet mask of Class-A network is 255.0.0.0, Class-B network is 255.255.0.0 and Class-C network is 255.255.255.0.

SELF-ASSESSMENT QUESTIONS - 3

6. IP addresses are represented by a _____ bit unsigned binary value.
7. The _____network 127.0.0.0 is defined as the loopback network.

INSPIRED BY LIFE

4. SUMMARY

- Protocols and types of protocols are used for communication between computers.
- Electronic mail, or E-Mail, lets you communicate with other people on the Internet.
- FTP uses TCP as a transport protocol to provide reliable end-to-end connections.
- Newsgroups started out as electronic bulletin board where people posted questions and answers.
- Discussion groups are similar to Newsgroups in the manner in which they work.
- The intenet security measures provides the protection on data over the internet.
- The domain name system (DNS) is the way that Internet domain names are located and translated into Internet Protocol addresses.
- A domain name is a meaningful and easy-to-remember handle for an Internet address.
- To identify a host on the Internet, each host is assigned an address, the *IP address*, or in some cases, the *Internet address*.
- IP addresses are 32-bit numbers represented in a *dotted decimal form*.
- The subnet mask is used by the internet layer to determine which part of the IP address is the network ID and which part is the host ID.

INSPIRED BY LIFE

5. TERMINAL QUESTIONS

1. Explain the services that are available for email.
2. Briefly explain the types of protocol.
3. What are the Internet security measures ?
4. Describe domain name system.
5. Discuss Class Based IP Address.

6. ANSWERS

Self Assessment Questions

1. E-mail or Electronic mail
2. File transfer protocol
3. C) Discussing group
4. iii. Firewall
5. secret key, password
6. 32-bit
7. Class A

Terminal Questions

1. E-Mail usually refers to a service that includes the store and forward, blind copies, offline delivery facilities. For more details refer section 3.1.2.
2. Transferring data from one host to another is one of the most frequently used operations. protocols specifies the transmission of data. For more details refer section 3.1.1.
3. The internet security measures are browser selection,multifactor authentication,email security and firewals.For more details refer section -3.2.2.
4. A domain name is a way by which a company can uniquely identify itself on the Internet. To understand the DNS and the way it is used, we need to understand the Internet naming structure. For more details refer section 3.3. .
5. To identify a host on the Internet, each host is assigned an address, the IP address, or in some cases, the Internet address. For classification of IP address refer section 3.3.1.

7. REFERENCES

- Internet Tim speed, Juanita Ellis (2003).*Security A Jumpstart for systems administrators and IT Managers.* Digital press.
- Charless Kozierok (2005).*TCP/IP Guide: A Comprehensive, Illustrated Interent* protocols.
- K. L. James (2003).*The Internet: A user's guide*. Second Edition.

INSPIRED BY LIFE