

Course Handout for Internet Technologies
(CSEU 08103) - Bachelor of Computer
Engineering

Amos Nungu, PhD
Dar Es Salaam Institute of Technology (DIT)

12.08.2013

Contents

1	Introduction	3
1.1	Introduction	3
1.1.1	Course Aim	3
1.1.2	Course Objectives	3
1.1.3	Course Outline	4
2	Internet Infrastructure and Architecture	6
2.1	Introduction	6
2.2	Chapter Objectives	6
2.3	History/Evolution of the Internet	6
2.4	Network Architecture	7
2.4.1	TCP/IP Protocol and The Internet Model	7
2.4.2	IP Addresses and Subnetting	9
2.4.3	Information Routing on the Internet	9
2.5	Other Key Internet Protocols, Standards and Infrastructures .	10
2.5.1	Standard Bodies	10
2.5.2	Domain Names	11
2.5.3	Domain Names and IP Addresses	11
2.6	Activity	12
3	Internet Applications and Services	13
3.1	Introduction	13
3.2	Objectives	13
3.3	Web Applications	14
3.3.1	Basic WEBAPPS Architecture	14
3.4	Common Web Applications	16
3.4.1	Blogs	16
3.4.2	The WiKi	16
3.4.3	Web Search Engine	17
3.4.4	Emails	17
3.5	Social and Commercial Issues of Web Applications	17

3.5.1	Internet Business Models	17
3.5.2	The Case of Social Networks)	18
3.5.3	Changing Economics of Information	20
3.5.4	Electronic Commerce	20
3.6	Emerging trends	22
3.6.1	Internet telephony	22
3.6.2	Streaming Multimedia Applications	22
3.7	Activity	22
4	Web Development	24
4.1	Introduction	24
4.2	Objectives	24
4.3	Developing Web Pages	24
4.3.1	Web Pages	24
4.3.2	Hypertext Markup Language (HTML)	25
4.3.3	PreExisting Software Packages and Services	26
4.3.4	Hosting	27
4.3.5	Programming Languages and Frameworks	27
4.3.6	HTTP Server Software	28
4.4	Integration of multimedia technologies	29
4.5	Activity	29
5	Security Concepts and Issues on Web	30
5.1	Introduction	30
5.2	Objectives	30
5.3	Security concepts and issues related to web technologies	30
5.4	Basic Cryptographic Concepts	31
5.4.1	Kinds of Cryptosystem	32
5.5	Cyber Security	32
5.5.1	Understanding Cybercrime	34
5.6	Activity	34

Chapter 1

Introduction

1.1 Introduction

Internet Technologies is a six(6) credit unit for Bachelor of Science (Computer Engineering) Degree course at the Dar Es Salaam Institute of Technology. The course discuss four major topics: Internet Architecture and Infrastructure, Internet Applications and Services, Developing Internet Applications, and Security Issues.

1.1.1 Course Aim

The major aim of the course is to provide you with an understanding of how the “Internet Works” as well as to expose you to practical application of the Internet in real life.

1.1.2 Course Objectives

Below are the comprehensive objectives of the course as stipulated in the curriculum. By meeting these objectives, you should have achieved the aim of the course as a whole. Upon Completion, you should be able to:

1. Present information using Internet Technologies
2. Secure information using Internet Technologies.
3. Manage information through the Internet
4. Determine the usage of computers in communication sector.
5. Search information.
6. Design Internet delivered applications.
7. Use Internet knowledge for electronic commerce.

8. Apply various approaches for authentication.
9. Identify the conflict between user-privacy and user tracking when using the Internet in electronic commerce systems

Prerequisite Module

CSEU 07402 - Multimedia Application Production

Proposed References per Curriculum

- 1 Greg Holden, (2002), "Starting on online Business for Dummies", John Wiley & Sons
- 2 Mark A. Miller, (2002), "Voice over IP Technologies, Building the converged network", John Wiley & Sons
- 3 Stephen Chen, (2001), "Strategic Management of e-business", John Wiley & Sons
- 4 Catherine Mambretti, (2003), "Internet Technology for Schools", Mc Farland & Company
- 5 Lawrence Harte, (2003), "Introduction to IP Telephony", Altos, ISBN.
- 6 Henry Sinn-Reich, (2004), "Internet Communications Using SIP", John Wiley & Sons
- 7 Intrigue Technologies Inc, (2002), "Intrigue Technologies Harmony Internet Remote", Intrigue Technologies Inc.

1.1.3 Course Outline

Based on the course objectives, this course is divided into four areas as follows:

Internet Infrastructure and Architecture

- History of the internet
- Network architecture
- The main internet protocols and standards
- Main Servers on the internet

Internet Applications and Services

- Structure of Internet/Web Applications
- Common Web Applications
- Social and Commercial Issues of Web Applications.
- Changing Economies of Information
- E-commerce

- Emerging Trends

Web Development

- Developing Web Pages/Application
- Database Driven Applications
- Integration of multimedia technologies.

Security Issues

- Security Concepts Related to the Web
- Some web Vulnerabilities
- Basic Cryptographic Concepts
- Cyber Security

Chapter 2

Internet Infrastructure and Architecture

2.1 Introduction

What is Internet? The Internet or simply the Net as it is often called, stands for interconnected network of networks on a global scale. It makes it possible for computers all over the world to send and receive messages. It is an internet-work of several hosts and their networks together to form a larger network of global magnitude. It is therefore a global collection of computers and networks that connects millions of peoples, organizations, military, and government to a wide range of information resources through a common protocol to communicate among themselves.

2.2 Chapter Objectives

At the end of this unit, you should be able to:

- Explain the Internet and its origin
- Explain the various application layer protocols
- Describe the concepts of domain names and IP addresses

2.3 History/Evolution of the Internet

The Internet came as a result of crave for a robust, efficient, store and forward, data network based on packets (packet switching) as against the circuit-switching (telephone network) previously in existence.

The packet switching technology was first implemented in the US Defense Advanced Research Projects Agency (DARPA) as ARPANET, a large area network developed by the Agency in late 60s. By 1986, the US Science Foundation (NSF) initiated a network called NSFNET, which later became a major component of the Net. Similarly, other networks were developed throughout the US through which the rest of the world was connected to form a global network of systems and network called the Internet [1, 2].

2.4 Network Architecture

Considering the global nature of the Internet, which involves several systems, running on different platforms, inter-system communication becomes an issue. This problem was solved using the transmission control protocol/Internet protocol (TCP/IP), which defines the rules for communication on the Net.

2.4.1 TCP/IP Protocol and The Internet Model

Transmission Control Protocol/Internet Protocol) is the basic communication language or protocol of the Internet.

The Internet Model predates the open standard interconnection (OSI) model. It dates back to the ARPANET (the origin of the Internet) and hence, often referred to as the Department of Defense (DoD) model.

The OSI model is composed of a seven-layered architecture along which network communications are segmented. Each layer covers a specific type of network activities, equipment or protocols. The layers are: Application, Presentation, Session, Transport, Network, Data Link, and Physical.

The DoD model is a four-layered architecture that does not map the OSI layer perfectly but can be relayed as indicated in Figure 2.1.

Layer 4: Process/Application Layer

This layer combines the functionality of the topmost three layers of the OSI model. That is, application, presentation and session. The protocols include those related to: e-mail, telnet, network management and directory services.

Layer 3: Host-to-Host Layer

This layer is the equivalent of the transport layer of the OSI model. It is responsible for end-to-end data integrity. The two protocols here are transmission control protocol (TCP) and the user datagram protocol (UDP). TCP

OSI		DoD	
Application Layer		Process/ Application Layer	
Presentation Layer			
Session Layer			
Transport Layer		Host-to-Host Layer	
Network Layer		Internetwork Layer	
Data-Link Layer		Network Access	
Physical Layer		Layer	

Figure 2.1: ISO and DoD Reference Model

offers reliable services and full duplex connections, while the UDP provides unreliable services that enhance throughput when error connection is not involved.

Layer 2: Internet Layer

This Layer corresponds to the network layer of the OSI model. Thus, it is responsible for routing packets within the internetwork. Gateways and routers are used for that purpose. The TCP/IP protocol at this layer is the Internet protocol (IP) which operates a system of logical host addresses called IP addresses.

Layer 1: Network Access Layer

This combines the functionalities of both the physical and data link layers of the OSI model. It is responsible for exchange of data between a host and the network as well as delivery of data between two devices on the same network. Communication between two systems through DoD model is as illustrated in Figure 2.2:

Communication between two systems over the Internet is described as follows: The packet of data is sent down the four layers of the DoD model from Application, through TCP, IP to Hardware (network cable, NIC) from where it is sent to the Internet and delivered to the destination system via the hardware (network cable, NIC), through IP, TCP and Application.

However, as the packet travels from one layer to another it wrapped into an envelope with the necessary information to the next layer. When received, it is unwrapped and rewrapped for the next lower layer. The procedure is repeated down the layers of the sending system and up the layers of the receiving system until it gets to the specified IP address. The process of

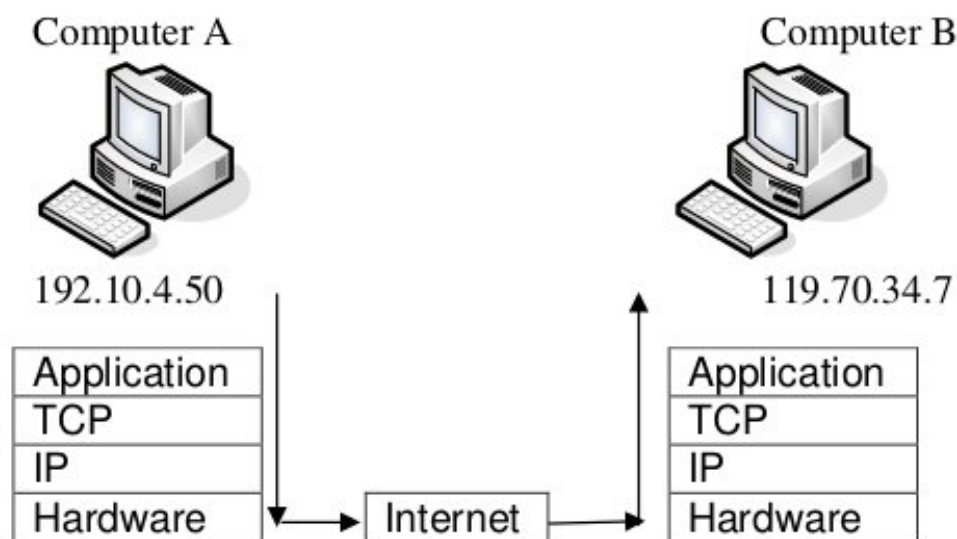


Figure 2.2: Packet Transmission through the Net

wrapping and unwrapping packets of data is also known as encapsulation and decapsulation respectively.

2.4.2 IP Addresses and Subnetting

IP addresses are unique four octet numbers expressed either as binary dotted or decimal dotted. e.g: 41.91.32.2

Address Classes

There are five different classes of address designed to meet the needs of different organizations. The various classes are given as A, B, C, D and E. The classes are distinguished from each other by the decimal notation of the first octet. The class range is presented in the Table 2.1:

Note: 127 (127.0.0.1) is reserved for loopback, for internal testing on the local machine. Web applications can be tested on the local machines using the loopback (localhost) before deployment to the web. Each system on a network or Internet is assigned a unique IP address by which it is identified.

2.4.3 Information Routing on the Internet

Data passage within an internet-work is via a router. That is, with the unique IP addresses assigned each of the systems, a packet of data is transmitted

Table 2.1: Different Classes of IP Addresses

Class	First Octet	Net-ID	Default Subnet mask	Availability
A	1-126	First Octet	255.0.0.0	Available
B	128-191	First 2 Octet	255.255.0.0	Available
C	192-223	First 3 Octet	255.255.255.0	Available
D	224-239	N/A		Reserved for multicasting
E	240-255	N/A		Reserved

from one node to another through the router.

The router maintains a routing table, which contains the IP address of all the adjacent nodes, the various subnets of the Internet. Similarly, it has the capability of routing packets of data through the shortest route using a number of routing algorithms. These make it a preferred candidate for a complex network of great magnitude like the Internet.

2.5 Other Key Internet Protocols, Standards and Infrastructures

2.5.1 Standard Bodies

A few bodies will be introduced here.

ICANN

The Internet Corporation for Assigned Names and Numbers (ICANN), has the responsibility of managing and assigning names and IP addresses. ICANN is a special organization with a charter from the US government to oversee the registration of names on the Internet alongside some domain name Registrars. More information could be accessed from their website <http://www.icann.org/> [Accessed in August 2013],

IETF

The Internet Engineering Task Force (IETF), has a goal of making the Internet work better. According their website, <http://www.ietf.org/> [Accessed in August 2013], their mission is to make the Internet work better by produc-

ing high quality, relevant technical documents that influence the way people design, use, and manage the Internet.

2.5.2 Domain Names

For easy identification and operations of the entire system (the Internet), the domain name service (DNS) maintains a database of computer names and IP addresses corresponding to each of them. That is, the DNS is responsible for translating domain names to IP addresses and vice versa as occasions demand between the users and the systems. Generally, domain names are composed of 3 or 4 parts namely: Hostname, Organization, Sub-domains (optional) and Country name.

e.g: Hostname.Organizationname.Type of categorization.Countrypname
OR

Hostname.Domain.Top-level domain.International top-level domain.

Where the top-level domains include:

- .com for commercial
- .mil for military
- .edu for education
- .org for organization
- .net for information services/network
- etc ...

and the international top-level domain use a 2-letter country code such as:

- .tz for Tanzania
- .us for United States
- .uk for United Kingdom
- etc...

e.g:

- www.openuniversity.ac.tz
- www.dit.ac.tz
- www.google.com
- www.cs.ac.uk

2.5.3 Domain Names and IP Addresses

For the sake of our study, it is important to understand that all the entities and inter-mediaries of the Net are given unique names for easy identification and delivery of messages. Thus, all the hosts/nodes on the Internet are identified through a unique identifier called domain names.

NOTE: Hosts are the organization who owns a particular network while nodes are the various computers connected the Internet via the host. The major parties involved in information dissemination are the users (people) and the systems (IT equipment). Therefore, domain names must be convenient and informative for all the parties to operate effectively. Users prefer the alphabetic naming conventions to recognize the hosts, while the computers or systems use the IP addresses.

2.6 Activity

1. Mention six network interconnecting devices.
2. Mention the device most appropriate for transmitting information over the Internet and why?
3. Mention five design goals of Internet Protocols
4. Describe the OSI model with reference to the DOD model.
5. What was the greatest weakness of store-and-forward networks?
6. Compare "packet" on the Internet to a POSTCARD
7. Relate an Internet router to a train station
8. What are Router Tables?
9. Compare a TCP port (such as port 80 for HTTP) to a Telephone extension
10. What is the value of a layered network model?
11. What is the Internet Protocol Layer responsible for?
12. Which is the lowest layer in the TCP/IP network model?
13. Mention 3 attribute of the Internet Layer
14. What does the Application Layer expect from the Transport Layer?
15. Is it possible to track a packet's journey across the network?, How?
16. What part of data transfer does TCP solve, and what part does IP solve?
17. In TCP, when does a sending system know it is safe to discard packets after it has sent them?
18. The storage of unacknowledged data is whose responsibility?
19. When is the Internet 100% up and working?
20. In the beginning there was Telephone networks, later came the Internet and last the convergence between the two. Compare and contrast a telephone network with Internet by mentioning four properties of each one.
21. Assuming you want to access an information somewhere from the Internet using your laptop; Mention, define and provide roles for all components required to assist you in accessing the content.

Chapter 3

Internet Applications and Services

3.1 Introduction

Most people these days depend on the World Wide Web to access the worlds information as well as connect them with people and events across the globe. These powerful online experiences are possible due to an open web that can be accessed by anyone through a web browser, on any Internet-connected device in the world.

The World Wide Web popularly called the “Web” is a global collection of interconnected documents on the Internet. Note that the web itself is not the Internet.

3.2 Objectives

At the end of this unit, you should be able to:

- Explain the structure of web application
- Identify some web servers available
- Search information.
- Explain the concepts of 3-tier and N-tier architecture
- Manage information through the Internet
- Social and Commercial Issues
- E-Commerce
- Emerging trends

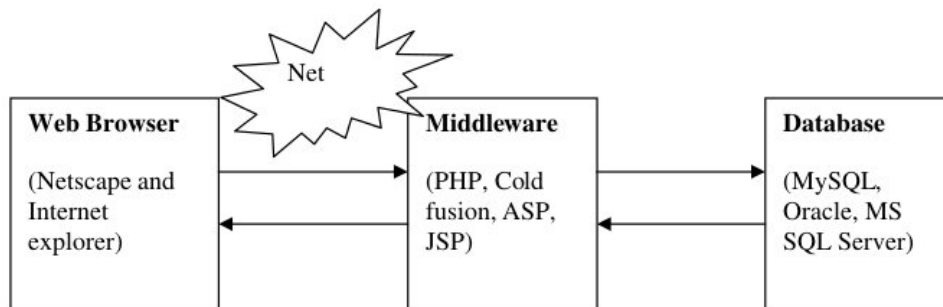


Figure 3.1: Structure of Web Application

3.3 Web Applications

Webapps are applications that are accessed with a web browser over a network such as Internet or Intranet. They are popular because of the ubiquity of the browser as a client (thin client). Similarly, the popularity is equally due to the possibility of updating and maintaining the application without necessarily distributing and installing it on every available client.

Webapps are used to implement services such as webmail, online retail sales, online auctions, discussion boards and weblogs among others. Web developers often use client-side scripting to add functionality to the webapps by creating an interactive site that does not require page reloading.

3.3.1 Basic WEBAPPS Architecture

The web works based on the Client/Server Architecture. That is both the server (web server or middleware) and the client (web browser) applications are responsible for some sort of processing. Web application is structured as a 3-tier application. That is web browser constitutes the first tier, a middleware engine using some dynamic web content technology such as: common gateway interface (CGI), hypertext preprocessor (PHP), java servlets or java server page (JSP), active server pages (ASP) constitute the middle-tier and the database being the third tier. Figure 3.1 depicts the structure of web application with example of tools at work in each tier.

The Web Browser (Client)

The web browser constitutes the client. It is a software application that enables a user to display and interact with text, images and other information. Browsers are used for quick and easy access to information contained in the web pages or at some websites by traversing some links. Similarly,

browsers can be used to access information on web servers. Examples of web browsers are MS Internet Explorer, Mozilla Firefox, Apple Safari, Netscape and Opera. Web browser works by communicating with web servers using the HTTP protocol. The protocol allows web browsers to submit information to web servers as well as fetch web pages from them. The primary language of browsers is the HTML, which consists of tags that are used to describe a web page. Most browsers also have some level of support for scripting languages like JavaScript and markup languages like extensible markup language (XML).

The Web Server (Middleware)

All Web transactions take place on the servers. The web server is responsible for communicating with the browser while the database server is responsible for storing the required information. The web server takes all requests from the clients, responds to the requests and serves the appropriate web pages back to the clients.

There are several web servers but the most prominent of them are Apache HTTP Server and the Microsofts Internet Information Services (IIS)

1. **Apache HTTP Server:** This is most popular web server. It is a free software/opensource. Apache runs on Unix, Linux, MS Windows, Novell Netware and some other platforms.
2. **Internet Information Services (IIS):** This is a major component of the Microsoft Server operating system, particularly, a component of its active server pages (ASPs). IIS is recommended if both the middleware (ASP) and the database Server (SQL Server) are Microsoft products. IIS is the worlds second most popular Web server after Apache.

The middleware is composed of languages such as PHP, ASP, ColdFusion, JSP and Perl. These languages work with web servers to interpret requests from clients, process the requests and interact with other programs that may be needed to fulfill the transactions and indicate to the web server the actual page to serve the client.

Database Server

This is a program that provides database services to other computer programs or computers. Database Management Systems (DBMS) provide functionality to database servers. They are responsible for storing, retrieving and manipulating the data in the database or other repositories. Some popular DBMSs include: Oracle, Sybase, Informix, SQL server, Db2 and Interbase.

3.4 Common Web Applications

All web applications communicate via protocols. In addition to the TCP/IP and DNS protocols, other common protocols include:

- HTTP This is known as hypertext transfer protocol. It is a connectionless text-based protocol through which clients (web browsers) send requests to the web server for certain web pages and images to be displayed. After servicing this request, the connection between the client and the server across the Net is disconnect. However, for another request from the client to be serviced, another connection is sought.
- Telnet This is an application layer protocol that enables users to execute terminal sessions with remote hosts. That is, it allows a user to login to another host at remote location.
- FTP This stands for file transfer protocol. It enables transfer of files between two hosts that are at remote locations to each other. It performs basic file transfer between hosts.
- MIME - Multi-purpose Internet Mail Extensions is an Internet standard that extends the format of email to support more than 7-bit ASCII characters.
- SMTP This stands for simple mail transfer protocol. This is a protocol that is used for exchanging electronic mail. It is used for basic message delivery.
- SNMP This is known as simple network management protocol. It is a protocol that is used to manage the network. It is used to collect information from the connected devices on the network for management purposes.

Common web application in our daily life include:

3.4.1 Blogs

Blogs are discussion or informational sites published on the World Wide Web and consisting of discrete entries (“posts”) typically displayed in reverse chronological order (the most recent post appears first).

In Tanzania, a good example is Michuzi blog available at <http://issamichuzi.blogspot.com/> [Accessed August 2013]

3.4.2 The WiKi

According to wikipedia [3], a wiki is a website which allows its users to add, modify, or delete its content via a web browser usually using a simplified markup language or a rich-text editor. Wikis are powered by wiki software.

Most are created collaboratively, Wikipedia is a good example. There are plenty of wiki software in the market, mostly open source.

3.4.3 Web Search Engine

A Web Search Engine is designed to search for information on the World Wide Web. Best example is google, and Bing. You can find an updated list of all web search engines at <http://www.philb.com/webse.htm> [Accessed August 2013].

3.4.4 Emails

Electronic mail, commonly referred to as email or e-mail, is a method of exchanging digital messages from an author to one or more recipients. Email types include:

- Web-based email (webmail) - browser is the client for accessing emails.
- POP3 (Post Office Protocol) email services - when accessing emails, they are downloaded from the server to the mail clients, no copy is left at the server - similar to Posta office.
- IMAP (Internet Message Access Protocol) email servers - when accessing emails, a copy is always left at the server until you explicitly delete it.

The protocol used for sending out email Simple Mail Transfer Protocol (SMTP).

3.5 Social and Commercial Issues of Web Applications

3.5.1 Internet Business Models

The following terms and/or definitions are associated with the Internet business models.

Dynamic pricing

Pricing of items based on real time interactions between buyers and sellers that determine what an item is worth at any particular moment.

Portal

Web site or other service that provides an initial point of entry to the web or to internal company data.

Banner ad

Graphic display on a web page used for advertising. The banner is linked to the advertisers web site so that a person clicking on it will be transported to the advertisers web site.

Syndicator

Business aggregating content or applications from multiple sources, packaging them for distribution, and reselling them to third-party web sites. Pure-play: Business model based solely on the internet. Clicks-and-mortar: Business model where the web site is an extension of a traditional bricks-and-mortar businesses.

3.5.2 The Case of Social Networks)

According to Wikipedia [4], “a social network” is a social structure made up of a set of social actors (such as individuals or organizations) and a set of the *dyadic* ties between these actors. Wikipedia also define “a social networking service” as a platform to build social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services.

A few examples of social networks include LinkedIn, Facebook, Twitter, Youtube, Google+, and Tumbra.

LinkedIn

LinkedIn is the professional network. Their Business Model is represented in Figure 3.2.

Facebook

Facebook is the dominant social network. Their Business Model is represented in Figure 3.3.

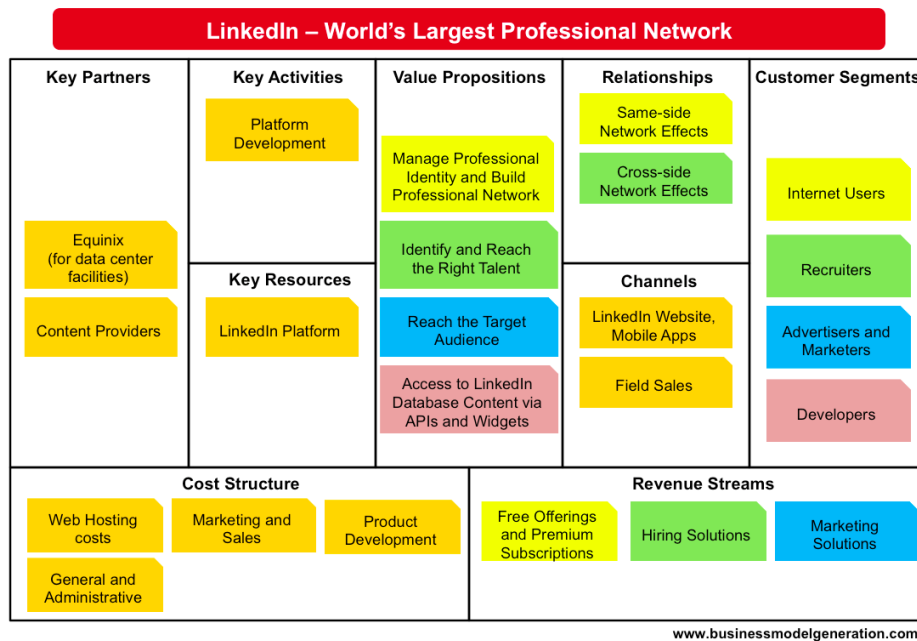


Figure 3.2: LinkedIn Business Model

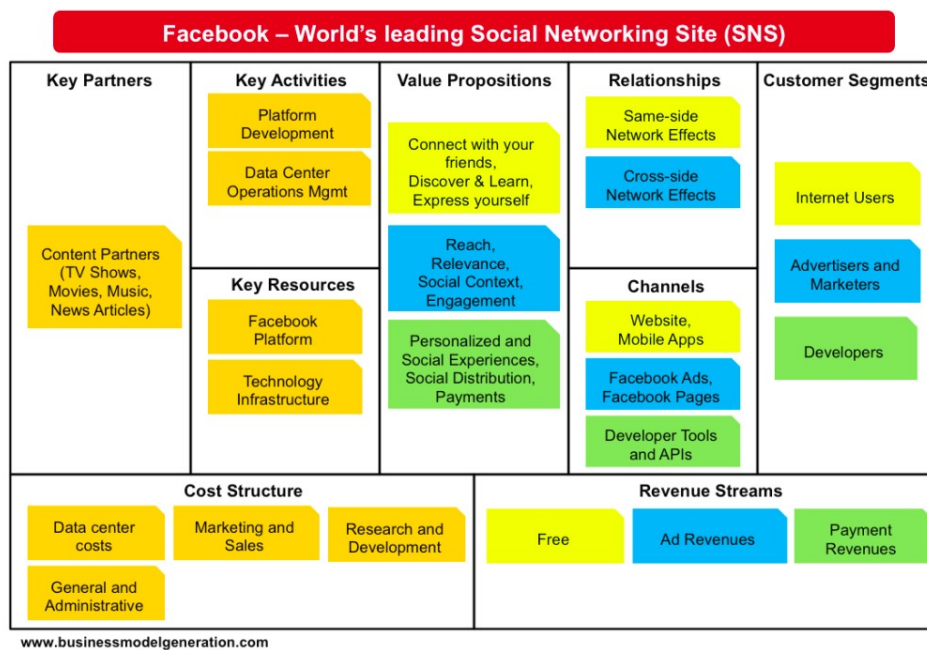


Figure 3.3: Facebook Business Model

Twitter

Twitter is the Micro-blogging platform. Their Business Model is represented in Figure 3.4

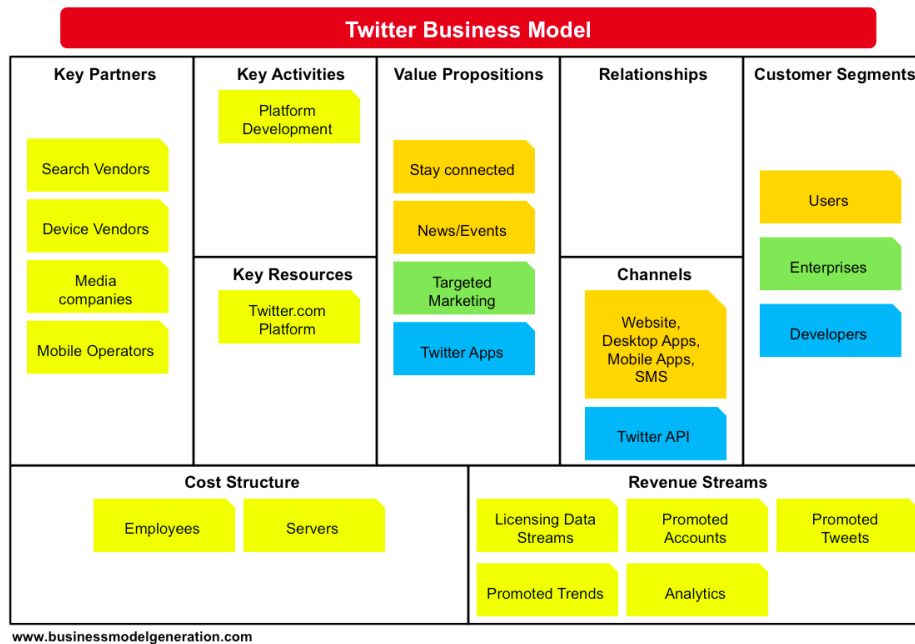


Figure 3.4: Twitter Business Model

3.5.3 Changing Economics of Information

- *Information Asymmetry*: Situation in which the relative bargaining power of two parties in a transaction is determined by one party possessing more information essential to the transaction than the other party.
- *Richness*: Measurement of the depth and detail of information that a business can supply to the customer as well as information the business collects about the customer.
- *Reach*: Measurement of how many people a business can connect with and how many products it can offer those people.

3.5.4 Electronic Commerce

Electronic commerce, commonly known as e-commerce or eCommerce, is a type of industry where the buying and selling of products or services is conducted over electronic systems such as the Internet and other computer networks. E-commerce draws on technologies such as mobile commerce, electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems. When shopping

online, the baskets for carrying goods are known as shopping carts.

Mobile commerce (m-commerce) is the use of wireless devices, such as cell phones or handheld digital information appliances, to conduct e-commerce transactions.

Types of e-commerce

1. *Business-to-Consumer(B2C)*: Electronic retailing of products and services directly to individual consumers.
2. *Business-to-business(B2B)*: Electronic sales of goods and services among businesses.
3. *Consumer-to-consumer(C2C)*: Consumers selling goods and services electronically to other consumers.

Payment systems for E-commerce

- **Digital wallet**: Software that stores credit card, electronic cash, owner identification, and address information and provides these data automatically during electronic commerce purchase transactions.
- **Micropayment**: payment for a very small sum of money, often \$1.00 or less.
- **Electronic cash (e-cash)**: Currency represented in electronic form that can be exchanged with another e-cash user or retailer over the Internet.
- **Smart card**: A credit card-size plastic card that stores digital information and that can be used for electronic payments in place of cash.

Cookies and Privacy in E-Commerce

Cookies are simple text files created for the main purpose of helping your browser process the special features of websites that use cookies. Cookies help website servers remember you as you navigate from page to page. This simple feature makes e-commerce possible since you don't have to reload your shopping cart every time you leave a page.

Cookies serve the following purposes:

1. Protection to make sure you are who you claim to be and not another person who managed to get a copy of your password.
2. Quickly determine your identity from page to page and remember the items you put in your shopping cart. This feature is essential for any type of e-commerce.

3. Settings cookies help the website you're visiting remember the settings you selected on a prior visit. This includes themes and language preferences as well as remember login names and passwords for easier entry on future visits.
4. Limit advertising cookies prevent "ad serving" scripts from showing annoying popup ads again and again. They also remember your previous pages so you don't see ads geared for first time visitors to those pages again during a session.

3.6 Emerging trends

We are going to discuss the Internet Telephony in this section.

3.6.1 Internet telephony

The term Internet telephony specifically refers to the provisioning of communications services (voice, fax, SMS, voice-messaging) over the public Internet, rather than via the public switched telephone network (PSTN).

Voice over Internet Protocol (VoIP) is a methodology and group of technologies for the delivery of voice communications and multimedia sessions over Internet Protocol (IP) networks, such as the Internet. Other terms commonly associated with VoIP are IP telephony, Internet telephony, voice over broadband (VoBB), broadband telephony, IP communications, and broadband phone service. VoIP is available on many smartphones, personal computers, and on Internet access devices. Calls and SMS text messages may be sent over 3G or Wi-Fi.

The Session Initiation Protocol (SIP), is the most used protocol in VoIP. Some common applications of VoIP include skype, viber and google talk.

3.6.2 Streaming Multimedia Applications

To be developed, the services include Video on Demand.

3.7 Activity

1. Mention 3 things Invented by Robert Cailliau and Tim Berners-Lee
2. What do you understand the term Servers? Give two example of a server
3. What is the purpose of the World-Wide-Web Consortium W3C?

4. What does the browser do when you click a Hypertext Link from your current web page to another web page?
5. Describe three types of e-commerce
6. Describe four payment systems for e-commerce
7. Mention any five social networks.
8. Define what is a business model and describe respective business models for networks mentioned above.

Chapter 4

Web Development

4.1 Introduction

The World Wide Web is a part of the Internet that uses the Hypertext Transfer Protocol (HTTP) to display hypertext and images in a graphical environment. It is a system of Internet servers that use HTTP to transfer documents formatted in HTML (Hypertext Mark-up Language).

4.2 Objectives

At the end of this unit, you should be able to:

- Explain Markup language for client side programming
- Describe Front-end and back-end scripting tools
- Design Internet delivered applications.
- Explain creation of hyperlinks and insert graphics on web pages
- Create forms and tables in web pages
- Design a simple static as well as dynamic web pages

4.3 Developing Web Pages

The chapter will discuss front end versus back end technologies, but also, server side scripting technologies and tools.

4.3.1 Web Pages

A Web page is a file with an .htm or .html extension. It contains HTML tags, but might also contain other code stored on a server. Basically, web pages are of two types: Static or Dynamic:

1. Static Web pages contains only HTML tags. They only display information and their content does not change when requested by the user.
2. Dynamic Web pages content changes depending on the user's request and preferences. Cannot be created using only markup language. It requires using Client-side code and Server-side (middleware) code. Dynamic web pages are also known as Database Driven websites.

4.3.2 Hypertext Markup Language (HTML)

This is a standard language for creating web pages. It provides tags that make documents look attractive using graphics, fonts and colours to enhance presentation. The basic building block of an HTML page is text, which are created using a Text Editor, while a Web Browser is used to test the application. The available text editors include Notepad or WordPad for Windows-based systems, while for Mac users, SimpleText is the HTML editor. UNIX users can use VI or Emacs.

Basically, an HTML document is composed of tags or elements. The tags are classified into container and empty tags. Coupled with HTML, the needed tools for web design can be described as follows:

Hypertext Transfer Protocol (HTTP)

This is a protocol for communication on the web and it supports the client-server model. The client-side communicates with the server-side through HTTP protocol.

Web Clients/browsers

The browsers interpret HTML documents. They are used to view texts, videos and audios as well as graphics files on web pages.

Hyperlinks

Hypertext links provide links to other portions of the web documents. When invoked, they forward requests to the server and the server responds with the appropriate response (document) to the client and it is displayed on the users screen.

Uniform Resource Locator (URL)

URL is an address/file specifier. It locates the documents requested through hyperlinks. Similarly, it may be used to point to a query, image or a command. The URL (Uniform Resource Locator) is the global address of documents and other resources on the web e.g. `http://www.who.int`. The first part of the address indicates which protocol to use e.g. `http`. The second part of the address identifies the domain name or the internet address where the information is located. Format:

Scheme://host-domain/path/dataname

`http://dit.ac.tz/downloads`

File Extensions used in Web Design

File extension	Data Type
Html/htm	html text
text/txt	ASCII textt
gif	composed graphicst
jpeg/jpg	composed graphicst
mpeg/mpg	digital videot
avi	digital videot
wav/au	digital video

4.3.3 PreExisting Software Packages and Services

Before deciding to create a web application from scratch, it is important to research existing software or services that are similar to what you want to build. This could include:

- software that you could use without any modifications
- opensource software that you could modify to fit your needs
- services where you could register as a customer

For example, if the primary purpose for your web application is publishing content on the web, then opensource content management system (CMS) software such as Wordpress or Drupal will typically be a good place to start. Or, you could use a hosted webpublishing service that requires no technical expertise, though at the expense of not allowing you to make changes to the software.

By researching existing software and services, you may find that you do not really need to develop your own software at all.

4.3.4 Hosting

All web applications must contain, at the very least, a server (also known as a host) located at some IP address that responds to requests over HTTP. An HTTP server can take a wide range of forms:

1. A regular laptop connected to a wireless or wired network
2. A dedicated desktop computer plugged into broadband internet in your office
3. A physical server that you rent inside a data center somewhere
4. A physical server where the operating system is shared with many other users
5. A virtual server where you have the operating system to yourself, but the physical machine is shared with other users
6. A virtual server that is automatically provisioned as necessary by a cloud provider

Each of these methods for setting up an HTTP server has different trade-offs. Criteria for evaluating hosting methods and providers include:

- Price: setup costs, monthly fees, bandwidth/CPU usage costs
- How reliable is the method in keeping your site online
- Support for your programming languages, software tools, and libraries
- Application performance / number of users that can be supported per server
- How easily you can scale to multiple servers when your traffic increases
- How easily can you switch to a different hosting method if necessary
- Network bandwidth and latency
- Setup time
- Required technical knowledge

The best choice for hosting your application will depend on your own needs and constraints.

4.3.5 Programming Languages and Frameworks

Typically, the architecture of web applications involves 3 basic layers, which typically use different programming languages:

1. A data storage layer, where commands are often expressed in SQL and executed on a server by database software such as MySQL, Microsoft SQL Server, or Oracle.
2. An application layer, which executes on a server and can use basically any programming language, including PHP, Python, Ruby, Perl, Java, JavaScript, C#, C++, Scala, Lisp, Haskell, etc. For many of these languages, there are multiple popular frameworks that make it easier

to develop web applications, such as Ruby on Rails (Ruby); ASP.NET (C#); Symfony, CakePHP, Yii, Zend, and Kohana (PHP); Django, Pylons, and TurboGears (Python); JSP, JSF, Struts, and Google Web Toolkit (Java); and Node.JS (JavaScript).

3. A presentation layer, which executes in a users web browser, and uses HTML, CSS, JavaScript, and occasionally ActionScript (i.e., Flash).

This section discusses the programming languages and frameworks used in the application layer, where there are a large number of different options to choose from.

The choice of application programming language is very important, because it is very difficult to change the language once you have written a large amount of code. There's no one best language or framework for every web application. There are many different criteria that you should examine when choosing a programming language to use, such as:

- your current experience in the language
- experience of other local developers, in case of support
- availability of documentation and tutorials
- readability/maintainability of code written in the language
- cost of developer tools and other required software
- speed of development (e.g. code verbosity, compilation time)
- compatibility with particular operating systems (Windows/Mac/Linux)
- availability of existing software libraries
- availability of compatible hosting providers
- performance / speed of applications written in that language

There are far too many programming languages and web application frameworks for anyone to be an expert in all of them.

4.3.6 HTTP Server Software

A HTTP server is a software program that listens for incoming HTTP requests from the network, and sends a corresponding response. The response may be a static file (e.g., an image) or output dynamically generated by your web application (typically HTML). There are numerous HTTP servers, such as Apache, Nginx, Lighttpd, and IIS, as well as many programminglanguage specific HTTP servers such as Tornado (Python), Tomcat (Java), and Node.JS (JavaScript). Depending on the HTTP server you use, your web application may run inside the HTTP server itself (perhaps via a module specific to your programming language), or in a separate process that communicates with the HTTP server via the CGI or FastCGI protocol. Criteria for choosing HTTP servers include:

- Compatibility with your operating system and programming language

- Availability of documentation and community support
- Ability to configure/customize to fit your needs
- Performance characteristics such as speed, CPU usage, and memory usage (which affect the number of simultaneous users each server can support)
- Reliability
- Price

4.4 Integration of multimedia technologies

To be expanded (Voice, video, flash, xHTML)

4.5 Activity

Assuming web design was covered in previous classes, every student should develop and host a database driven website that can store and display the following contents: Your Name and Registration Number.

Chapter 5

Security Concepts and Issues on Web

5.1 Introduction

Security is always a Trade-off, “Perfect security” is unachievable. One must find the right trade-off between Security versus Cost or Security versus Convenience.

5.2 Objectives

At the end of this unit, you should be able to:

- Secure information using Internet Technologies.
- Apply various approaches for authentication.
- relate security issues including symmetric and asymmetric key, encryption and digital signature, authentication.

5.3 Security concepts and issues related to web technologies

This section will explain some common Web application vulnerabilities include: Remote code execution; SQL injection; Cross Site Scripting (XSS); and Username enumeration [7, 6, 8, 9].

Remote code execution

This vulnerability allows an attacker to run arbitrary, system level code on the vulnerable server and retrieve any desired information contained therein. Improper coding errors lead to this vulnerability.

SQL Injection

Serious Websites are Database Driven, hence dynamic. SQL injection technique allows an attacker to retrieve crucial information from a Web server's database. Depending on the application's security measures, the impact of this attack can vary from basic information disclosure to remote code execution and total system compromise.

Cross Site Scripting

Cross Site Scripting is an attack in which a user is tricked into executing code from an attackers site (say evil.com) in the context of legitimate site. When visiting such a crafted URL, an attacker can effectively execute something malicious in the victim's browser. Some malicious Javascript, for example, will be run in the context of the web site which possesses the XSS bug.

Cross-Site Request Forgery (CSRF)

Cross-Site Request Forgery (CSRF) is an attack where a malicious site tricks our visitors into carrying out an action on our site. This can happen if a user logs in to a site that they use a lot (e.g. Facebook, etc.), and then visits a malicious site without first logging out

Username Enumeration

Username enumeration is a type of attack where the backend validation script tells the attacker if the supplied username is correct or not. Exploiting this vulnerability helps the attacker to experiment with different usernames and determine valid ones with the help of different error messages (e.g. Valid username with incorrect password).

5.4 Basic Cryptographic Concepts

This section will look into Cryptographic concepts including Cryptographic Hash function, symmetric and asymmetric key, encryption and digital signature, authentication and authorization.

Some basic terminologies:

1. **Cryptographic Hash function** - takes a block of data and computes a fixed-size bit string called the hash value
2. **Confidentiality** - Prevent unauthorized viewing of private information
3. **Integrity** - Information is from who you think it is from and has not been modified since it was sent
4. **Plaintext** - is a message that will be put into secret form
5. **Ciphertext** - is a transformed version of plaintext that is unintelligible to anyone without the means to decrypt
6. **Encryption** - The transformation of plaintext to ciphertext
7. **Decryption** - Returning the ciphertext back to plaintext

5.4.1 Kinds of Cryptosystem

The strength of a cryptosystem is determined by the encryption and decryption techniques and the length of the key.

1. Two basic types of cryptosystems exist, secret-key and public-key.
2. In a secret-key scheme, the key used for encryption must be the same key used for decryption. Also called symmetric-key cryptosystem.
3. Secret-key cryptosystems have the problem of secure key distribution to all parties using the cryptosystem.

Digital Certificates

In cryptography, a public key certificate (also known as a digital certificate or identity certificate) is an electronic document which uses a digital signature to bind a public key with an identity information such as the name of a person or an organization, their address, and so forth. The certificate can be used to verify that a public key belongs to an individual.

Certificate Authority (CA)

A certificate authority is an entity that issues digital certificates. The digital certificate certifies the ownership of a public key by the named subject of the certificate. A CA is a trusted third party that is trusted by both the owner of the certificate and the party relying upon the certificate.

5.5 Cyber Security

Some definitions and explanations:

- The “Information and Communication Technologies” or **ICT** refers to “technologies that provide access to information through telecommunications. It is similar to Information Technology (IT), but focuses primarily on communication technologies. This includes the Internet, wireless networks, cell phones, and other communication mediums” [5].
- **Cybersecurity**: According to the ITU Framework, cyber security is defined as: the prevention of damage to, unauthorized use of, exploitation of, and – if needed – the restoration of electronic information and communications systems, and the information they contain, in order to strengthen the confidentiality, integrity and availability of these systems.
- **Cyber Crime**: Criminal acts using computers and networks as tools or targets Traditional crimes conducted through the use of computers
- **Security Attributes**: Three Security attributes or qualities are: Confidentiality, Integrity and Availability (CIA).
 1. *Confidentiality* - prevent the disclosure of information to unauthorized individuals or systems
 2. *Integrity* means that data cannot be modified undetectably.
 3. Information must be *Available* when it is needed
- **Risk Management**: Is the process of identifying vulnerabilities and threats to the information resources used by an organization in achieving business objectives, and deciding what countermeasures, if any, to take in reducing risk to an acceptable level, based on the value of the information resource to the organization CISA Review Manual 2006.
- **Risk Components of ICT**:
 1. People (The weakest point in the chain)
 2. Procedures (such as hiring, promoting, and outsourcing)
 3. Software (such as patches and upgrading)
 4. Hardware (such as configuration errors and patches of software (firmware) bugs)
- **Common Threats/Risks**: Categorized into malicious codes, network attacks and network abuses.
 1. *Malicious codes* include computer viruses, worms, Trojan horses, spyware, key loggers, and BOTs.
 2. *Network attacks* include intrusions, DoS (Denial of Service) attacks and web defacement.
 3. *Network abuse* includes SPAM, phishing, pharming and network-related forgery.

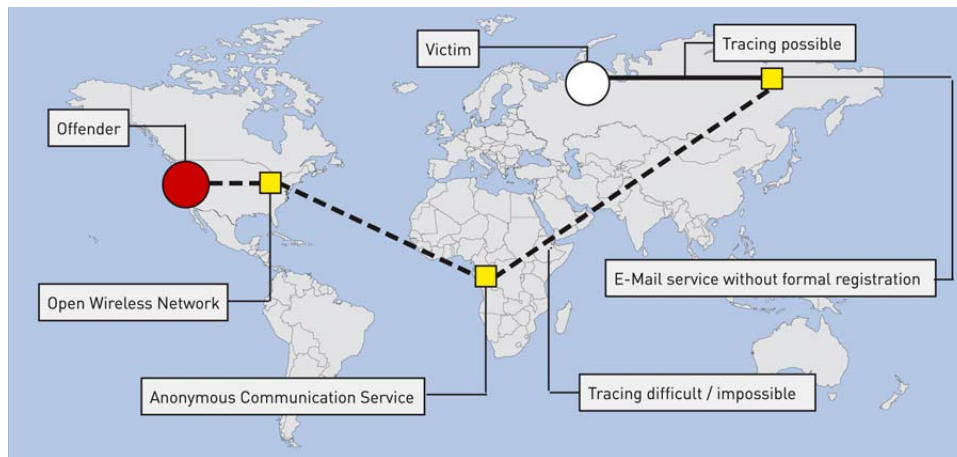


Figure 5.1: ITU-D: Cybercrime Legislation Resources 2011

5.5.1 Understanding Cybercrime

The diagram in Figure 5.1 represents a Transnational Dimension of the Cyber Crime.

The Problem

As stated by Bruce Schneier, April 2009: “*Cybersecurity isn’t a military problem, or even a government problem – it’s a universal problem. All networks, military, government, civilian and commercial, use the same computers, the same networking hardware, the same Internet protocols and the same software packages. We all are the targets of the same attack tools and tactics*”.

5.6 Activity

1. When you are using secure http and sending data between your computer and your bank’s computer, where is the data encrypted and decrypted?
2. Considering the four-layer TCP/IP model of Application / TCP / IP / Link, which two layers does Secure Sockets Layer (SSL) fit between?
3. What does a cryptographic hash function do?
4. If you are sending credit card information from a coffee shop WiFi to an Internet web site and later you find your credit card information has been stolen, which is the most likely scenario as to how your information was stolen?
5. In regards to security, what do we assume about the Internet?

6. When is Integrity preserved?
7. What is packet sniffing?
8. Mention and explain any five online security threats
9. In security, what is the relationship between SSL and TLS?

Bibliography

- [1] History of the Internet - <http://www.isoc.org/internet/history/> [accessed on August 2013]
- [2] Internet chronology 1960 - 2001 - <http://www.packet.cc/internet.html> [accessed on August 2013]
- [3] Wikipedia - <http://en.wikipedia.org/wiki/Wiki> [accessed on August 2013]
- [4] Social Network - http://en.wikipedia.org/wiki/Social_network [accessed on August 2013]
- [5] The Tech Terms Computer Dictionary - <http://www.techterms.com/definition/ict> [accessed on August 2013]
- [6] <http://coding.smashingmagazine.com/2010/10/18/common-security-mistakes-in-web-applications/> [Accessed August 2013]
- [7] https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project [Accessed August 2013]
- [8] <http://www.symantec.com/connect/articles/five-common-web-application-vulnerabilities> [Accessed August 2013]
- [9] http://sharecase.ucsd.edu/_files/WAS-Sharecase2012.pdf [Accessed August 2013]