زبان تخصصی
مهندسی کامپیوتر

پوپول مرجع دانشگاه و مدرسه
www.pupuol.com

TECHNICAL AND VOCATIONAL UNIVERSITY

English for students of computer science

Tabriz Junior Collage

By: Sattar Hashemi Khosroshahi

2012

IRAN - TABRIZ

Contents

Lesson one

UNDERSTANDING THE OSI REFERENCE MODEL[1]

In 1984, the International Organization for Standardization (ISO) and what is now the Telecommunication Standardization Sector of the International Telecommunications Union(ITU-T) published a document that divides the functions of a data network into seven layers,as shown in Figure 1. "The Basic Reference Model for Open Systems Interconnection,"now commonly known as the OSI reference model (ISO/IEC 7498-1:1994 and ITU-T Recommendation X.200), has become an industry standard for teaching and referring to networking functions.

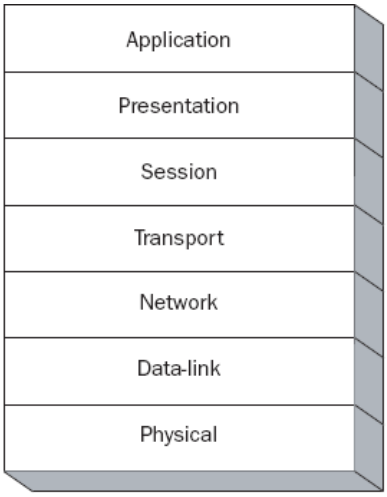


Figure 1 : The OSI reference model

The seven layers of the OSI model define functions that are implemented in various networking protocols, such as Ethernet and TCP/IP. The functions corresponding to the layers are as follows:

■ ***Physical*** The physical layer defines the nature of the network medium—the actual fabric of the network that joins all the computers together—and the nature of the signals transmitted over the medium. In most cases, the network medium is a form of copper cable that uses electric currents for signaling, but fiber-optic andwireless media are becoming increasingly prevalent.

■ ***Data-Link*** The data-link layer defines the interface between the network medium and the software running on the computer. Among the data-link layer functions are packet addressing (which allows computers to direct traffic to specific destinations on the local network); media access control (which allows multiple computers to share a single network medium without conflicting); and formatting the frame used to encapsulate data for transmission over the network.

The data-link layer divides into two sublayers. The logical link control (LLC) sublayer controls elements such as error checking from node to node on the same LAN, frame synchronization, and flow control. The media access control (MAC) sublayer controls the movement of data packets to and from one network interface card (NIC) to another across a shared channel.

■ ***Network*** The network layer defines the functions that provide end-to-end communications between computers on different networks. Chief among these functions is routing, which enables

computers to relay traffic through intermediate networks to a destination on a remote network any distance away. Other functions include packet sequencing, end-to-end error detection from the sender to the recipient, congestion control, and addressing. While the data-link layer is responsible for local traffic on a single network, the network layer is responsible for directing traffic to its ultimate destination.

■ ***Transport*** The transport layer provides functions that complement those of the network layer, including guaranteed delivery (which uses packet acknowledgments to ensure data is received), flow control (which regulates transmission speed to avoid dropped packets), and end-to-end error detection (which enables the receiving system to detect damaged packets).

■ ***Session*** The session layer provides many functions involved in the regulation of the dialog between two computers communicating over the network. For example, the session layer sets up, regulates, and terminates exchanges between the applications at each end of the communication.

■ ***Presentation*** The presentation layer (sometimes referred to as the syntax layer) is responsible for translating each computer's native syntax into a common transfer syntax readable by the other computers on the network. In some cases, the transfer syntax can provide functions such as data compression and encryption.

■ ***Application*** The application layer provides the interface between the networking protocol stack and the software running on the computer. For example, this layer provides the interface for e-mail, file transfers, Telnet and File Transfer Protocol (FTP) applications. Applications use the services provided by applicationlayer protocols, which in turn use the services provided by the other layers beneath them.It is important to understand that the protocols that implement the functions of the OSI model do not correspond exactly to the individual layers. A computer on a network does not necessarily run seven different protocols, with one corresponding to each layer. Generally speaking, the designer of a network infrastructure selects a data-link layer protocol, such as Ethernet or Token Ring, which actually encompasses both the physical and data-link layers in its functions, and a protocol suite, such as TCP/IP,which implements the functions of the network and transport layers. The session, presentation,and application-layer functions are sometimes provided by a protocol in thesuite or by a separate application-layer protocol.

Selecting a Data-Link Layer Protocol

The selection of a data-link layer protocol is the most important decision in the design of the network's physical infrastructure. The data-link layer protocol is not only responsible for strictly data-link layer functions, such as media access control, but also for the network's physical layer implementation. Currently, the most commonly used data-link layer protocol on networks is Ethernet, with Token Ring running a distant second. However, there are several Ethernet variations that provide various levels of performance, and selecting the correct one is crucial.

You need to consider a number of criteria when selecting the data-link layer protocol for use on a network. Because the data-link layer protocol dictates the nature of the network's physical infrastructure, you must consider design elements such as the distance between workstations and the transmission speed you require. You must also consider the nature of the traffic the network will carry and its amount. Additionally, your budget is always an important consideration.

Selecting a Media Type

Although there have been other types of media used in the past, most LANs constructed today use either unshielded twisted pair (UTP) cable or fiber-optic cable. In most cases, UTP cable is sufficient and much less expensive, but fiber-optic cable provides a viable alternative when you have special performance requirements. Network media that consist of cables are called *bounded media*. Recent developments in wireless LAN technologies have also made *unbounded media* (networks that don't use cables) a viable and practical solution.

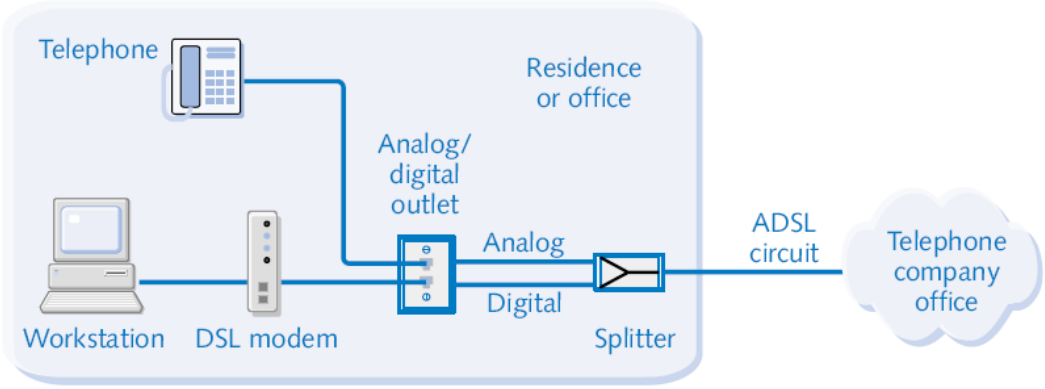
Lesson two

DIGITAL SUBSCRIBER LINE (DSL) CONNECTIONS

DSL is a broadband technology that has recently gained tremendous popularity. To use DSL, your computer must be outfitted with a *DSL modem*. Some DSL modems are external devices that connect to a computer's NIC or to a hub, switch, or router via an RJ-45 connector (or directly to a computer via a universal serial bus [USB] port). Other DSL modems are internal, plugging directly into an expansion slot on acomputer's motherboard.

Similar to a dial-up modem, DSL also works over public telephone lines (but only those where the DSL service is available). DSL uses different channels to transmit high-speed digital data over the phone company's wiring. Telephone lines are made up of pairs of copper wires, and the pair that transfers analog voice data can also be used simultaneously to transfer high-speed digital data on different channels. The channels are effectively merged into one signal in the copper wire. The merging is accomplished by conversion equipment at the phone company's central office, and the signal is later separated at your home or office by a splitter or by using filters to separate the analog anddigital signals.

Because specialized equipment is required in the central office that serves your home or office, the phone company must have this equipment in place to support DSL in your area. In this way you can use your phone normally while also using the DSL modem on the same phone line. In other words, DSL uses your phone line's existing pair of wires—no additional lines are required, as shown in the following illustration.



Most DSL data services use asymmetric DSL (ADSL), which divides the phone line into three channels. One channel is used for voice, a second channel is used for data transmission from your computer (including Web page requests and files or e-mail you send), and a third channel is used to transfer data back to your computer. The throughput is slower for the channel used to request and send information, whereas the receiving channel provides greater throughput because in most situations moredata is downloaded than uploaded.

Some DSL providers offer symmetric DSL (SDSL). Unlike ADSL, SDSL uses the same speed on both its uplink and downlink channels. While most DSL providers today offer ADSL because most

personal users do not upload large amounts of data and simply do not need as much upstream bandwidth, if you need to transfer large files often or host an Internet server, SDSL is the better choice.As a general rule, DSL installation is easy, works with most computers, and is very reliable.
Many providers will even give you a free DSL modem if you sign up for at least a year of service. DSL service typically costs around $40–$60 a month, and it will work with Windows XP Internet Connection Sharing (ICS) or other Internet connection sharing software, such as WinProxy. Because the DSL connection works over the phone ine, the connection belongs to you alone (there is no sharing with other users as with cable Internet access), and the connection is always available—there is no phone numberto dial (although you might periodically have to re-establish a client PPPoE connection—for more information, see "What Is PPPoE?" on page 105). Data transfer occurs automatically over the line to the DSL provider.
Sound too good to be true? Well, for many people it is. The biggest problem with DSL is that you must have a DSL provider in your area, and you must live within a certain distance of that provider's offices. DSL traffic can only travel a certain distance before degrading, so even though your phone company provides DSL, you still might not qualify, depending on how far away you live. Additionally, although DSL connections are not shared, they are often bandwidth-limited, requiring that you pay a higher monthly cost if you want to match the maximum throughput of cable Internet connections. Table 4-2 summarizes the characteristics of DSL Internet.

What Is PPPoE?

You might have noticed the reference to Point-to-Point Protocol over Ethernet (PPPoE) when you selected the type of connection that you wanted to create. *PPPoE* is a type of broadband Internet connection that is not always connected, but instead requires a user name and password to be sent each time the user wants to connect.
PPPoE is designed for users on a LAN (using standard Ethernet) who access the Internet over an Ethernet network through a broadband connection. In other words, Point-to-Point Protocol (PPP), which is used on the Internet, functions over Ethernet to provide Internet access to these users. With PPPoE, each user can have a different access configuration, even though they all reside on the same LAN.

Lesson Three

INTRODUCTION TO THE .NET FRAMEWORK

- ✓ The .NET Framework is a platform for building and running applications.
- ✓ A platform for building, deploying, and running web services and applications.
- ✓ Microsoft .NET provides prefabricated infrastructure for solving the common problems of writing Internet software.

In addition, .NET enables you to write programs or applications for a distributed environment. To do so, .NET helps you to create Web services and Web applications. The .NET Framework not only helps you write new programs, but also provides you with the ability to improve the existing programs. The .NET Framework is well designed for communicating with existing COM (*Component Object Model*) components, making the applications written in .NET languages backward compatible with existing programs.

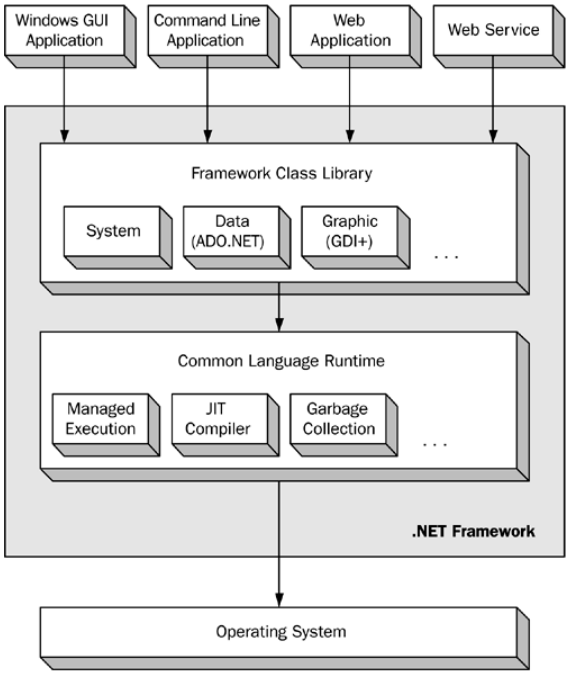


Figure 1.1. The .NET Framework

The .NET Framework provides a complete development framework. The .NET Framework helps you write complex applications by providing you with predefined classes and methods in the .NET base class library and also manages the execution of the applications that you write.

.NET provides you with a library of classes, which contains several base classes called *.NET base classes*. These classes, in turn, define several functions that you can use to write your applications. In addition,.NET provides the .NET Runtime environment called the CLR to execute the code written for the Windows platform.

.NET offers an application development environment called *Visual Studio .NET* that consists of several programming languages, such as Visual Basic .NET, Visual C#, Visual FoxPro, and Visual

C++ .NET. These programming languages combine the features of the existing languages with several new features to provide a powerful development framework. Following are some of the features of the .NET Framework.

◆**Interoperability with other environments.** The need for a new development environment was primarily because the applications developed in existing environments were not platform-independent. For example,applications that you develop for the Windows platform are not compatible with the applications designed for the UNIX environment.With the evolution of the .NET Framework, you can develop applications that can run on the Internet, making them accessible across various platforms. These applications are called *Web applications* and are fully supported by the .NET Framework.

Interoperability across environments, which is the strongest feature of .NET applications, is a result of .NET's support for MSIL (*Microsoftintermediate language*). At the time of compilation, all the managed code written for the .NET platform is converted to MSIL, which is a set of CPU-independent instructions. When you run the code written for the Windows platform in any other environment, for example UNIX, the compiler compiles the MSIL code to one that UNIX understands,enabling the application to run.

◆ **Support for developing language-independent applications.**

In addition to developing applications that can interoperate with those in a different environment, you can develop applications that are language independent.Visual Studio .NET provides a common development environment for all languages in the .NET series. This implies that if you develop an application by using any language of the .NET family, the code can be easily translated and used by any other .NET language. For example, a Visual C++ .NET application can be easily converted to a Visual Basic .NET or Visual C# application, and likewise the opposite.

◆ **Support for OOPs (*object-oriented programming*).**

OOPs is not a new concept for C++ programmers. Code in a OOP-based language is written using classes and objects. This not only helps you to write code easily,but also helps reuse your code. As discussed earlier, .NET has a library of classes that contains methods that you can use to develop your applications. In addition, .NET also supports inheritance, which means that you can derive new classes from existing or base classes, and thus make the base class methods available to the new classes.

◆ **Support for Web applications.**

CreatingWeb pages using scripting languages such as ASP (*Active Server Pages*) has not been an easy task for programmers worldwide. Therefore, to make coding simpler for the programmers, .NET provides the ASP.NET technology. The applications that use the ASP.NET technology to create Web pages are called Web applications.Using ASP.NET, you can create new ASP.NET pages or convert existing ASP pages to ASP.NET pages. ASP.NET also enables you to add highlevel functionality to your Web pages by allowing you to create pages in any of the .NET programming languages. For example, using ASP.NET,you can create dynamic Web pages that allow you to access data from an underlying database.

◆ **Support for Web services.**

.NET helps you to create Web services that you can use to create applications for different platforms that access data through the Internet. To do so, the methods of an instance of a class are called across the Internet and can then be used by applications running on various platforms. In addition, Web services help you to access the functionality of a remote server, such as calling a method from a remote server, creating an instance of a class on a remote server, and performing operations on the remote server.Web services use HTTP, which simplifies your task of accessing a remote server. HTTP helps in transferring messages written using XML between client and server.

Common Language Runtime (CLR)

Among the most important components of .NET is the ***.NET Runtime***, commonly called the **CLR**. As the name suggests, the CLR is a common run-time environment for the code written in .NET languages.The code in .NET is managed by the CLR and is, therefore, called *managed code*. Managed code contains the information about the code, such as the classes, methods, and variables defined in the code. This information contained in managed code is called *metadata*.

The CLR uses metadata to provide safe execution of the program code. In addition to executing code, the CLR manages memory and threads and helps in the security and interoperability of the code with other languages. Besides providing safe execution of the program, managed code aims at targeting CLR services. These CLR services include locating and loading classes and interoperating with the existing DLL (*Dynamic Link Library*) code and COM objects.The CLR has also enabled programmers to achieve interoperability across applications written in any of the .NET languages. Because the CLR is the common runtime environment for all .NET applications, all the code in .NET is converted to MSIL and is executed in a similar fashion. As discussed earlier, this code is called managed code. Managed code in .NET is developed using the CTS or CLS classes.

Common Type System (CTS)

As discussed earlier, .NET aims at providing interoperability between applications.To create interoperable applications, you need a set of standard data types that would be used across applications. In addition, you require a set of guidelines to create user-defined classes and objects for the .NET Framework. These standard data types and the set of guidelines are contained in CTS. To ensure interoperability across applications, CTS includes only those data types and features that are compatible across languages.Consider an example of an applica t i on of which a part is created using C++. Subsequently, to provide a visual interface, you need to recreate the entire application in Visual Basic. This means that you need to re c reate all the classes that you have used in the C++ application . This is because C++ and Visual Basic are not interoperable .The CTS feature for the .NET Framework simplifies such tedious tasks.If you create a class in any of the languages in the .NET Framework , you can use the same class in another language that is supported by the .NET Framework .

Common Language Specification (CLS)

In addition to CTS, another feature that ensures language interoperability in the .NET Framework is CLS. CLS is defined as a set of rules that a .NET language should follow to allow you to create applications that are interoperable with other languages. However, to achieve interoperability across languages, you can only use objects with features listed in the CLS.

Lesson Four

THE REVERSE ENGINEERING PROCESS AND TOOLS

Reverse engineering is defined by Chikosfky and Cross as the process of analyzing a subject system to identify the system's components and their relationships, and to create representations of the system in another form or at a higher level of abstraction. The process of reverse engineering is accomplished using specific tools that, for the 32-bit Microsoft Windows environment, are categorized as hex editors, disassemblers/debuggers, decompilers, or related technologies such as code obfuscators, unpackers, and PE editors. An evaluation of each tool is provided that identifies its domain of applicability and usability.

The Reverse Engineering Process

Software engineers are sometimes asked to understand the behavior of a program given that program's binary executable file. If they have access to the appropriate reverse engineering tools, they might choose to adhere to the following process.First, a general disassembler/debugger is used to determine the basic functionality of the program. If disassembly and debugging shows that the binary code has been obfuscated, the next step would be to determine whether the obfuscator used is a common commercial obfuscator or a custom protection scheme. A PE editor would be used to make this determination. If the obfuscator used was a common commercial obfuscator, an unpacker could be used to restore the original code. If the obfuscation has been customized, however, manual unpacking would be necessary.

This is accomplished by tracing the execution of the program in a debugger until the original code is found. Then, a memory dumper can be used to write the program to the disk, and a PE editor can be used to restore the PE Headers and make the program executable again.

Once the obfuscation has been cracked, or if there was no obfuscation used, the process continues with a disassembler/debugger. Typically, a debugger can be used to locate code that needs to be changed, such as text strings, message box calls, or read/write functionality. To alter this code permanently, an opcode patch needs to be applied in a hex editor. Once the change has been made and applied using a hex editor, an executable will be created with the new functionality.

In some cases, using a disassembler/debugger may reveal that the code being examined was written in Delphi, Java, or Visual Basic. In the case of Java, the next step would be to use a decompiler to examine the code. In the case that the code is written in Delphi or Visual Basic, which is often made evident from the headers and the strings, a disassembler that is targeted speci_cally toward one of these two languages should be used. These special disassemblers can provide much more analyses of the code, and a more accurate disassembly. These programs also provide a convenient interface to alter executable files and save changes to them onto the disk.

Often, when the end of the aforementioned reverse engineering process is reached, a software engineer must return to the beginning of the process to evaluate the outcome of this process and possibly make additional changes. Sometimes, the effectiveness of these changes is compromised by anti-reverse engineering techniques.These anti-reverse engineering techniques include debugger detection, which attempts to prevent the use of reverse engineering tools, and checksums, which are computed values used to check whether changes have been made to the executable files of programs. Often, multiple iterations of this process are necessary to defeat anti-reverse engineering

code, eliminate checksum checks, and so on. Figure 1 shows a flow chart of the reverse engineering process.

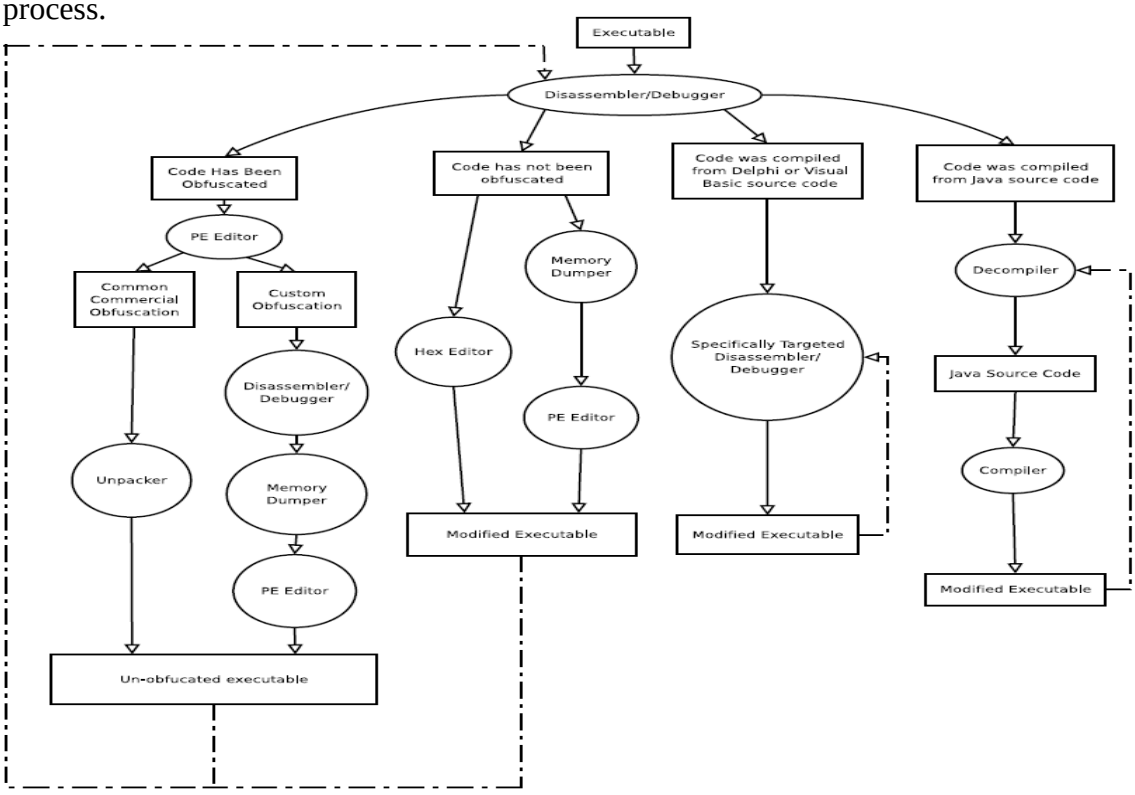


Fig. 1. The Reverse Engineering Process

Categorization of Reverse Engineering Tools

Since the determination of target functionality of a binary executable is hardly practical, an arsenal of tools has been developed over the years to assist in data gathering, extraction, organization, and classification. In this survey, we concentrate on tools, which we describe in four categories, for the Microsoft Windows environment.

(1) Hex editors are programs that facilitate the editing of binary files using, typically,a hexadecimal (base-16) representation of the binary data. Hex editors are also able to display ASCII and Unicode equivalents of the hexadecimal format.

(2) Disassembler/Debuggers are programs that support the translation of hexadecimal text into assembly language, which, although human readable, is often not as intuitive as the original source code. This translation is performed by disassemblers.Debuggers can stop a program's execution at specified locations in the code to examine the program's state, thereby assisting in the comprehension of executing programs. Disassemblers and debuggers are often combined into a single tool that displays a program in assembly language and allows the execution of the program to be controlled.

(3) Decompilers attempt to translate an executable program into source code. Usually specific to a particular compiler, decompilers also display assembly language code for parts of the program they were unable to decompile.

(4) Related technologies such as code obfuscators, PE editors, memory dumpers, and unpackers have specific applications in reverse engineering. Code obfuscators can operate on either source code or binary code in order to make the code more dificult to understand. PE editors extract the headers from PE files and allow them to be edited more eficiently than by using a hex editor.

Memory dumpers allow a debugged program that has been altered in memory to be saved to disk. Unpackers are designed to overcome commercial protection techniques.

Lesson Five

WHAT IS DATA MINING?

Data mining is the task of discovering useful knowledge automatically from large data repositories. The notion of usefulness has different meanings to different people.From a business perspective, the knowledge is useful if it helps managers to make the right business decisions. For Earth Scientists, the knowledge is useful if it reveals previously unknown information about how the earth system is changing over time.Data mining is often considered to be an integral part of knowledge discovery in databases (KDD). KDD is the overall process of converting raw data into useful knowledge, as shown in Figure 1. It comprises of a series of transformation steps, including data preprocessing and postprocessing. The goal of data preprocessing is to transform the raw data into a suitable format for subsequent analysis. It also helps to identify attributes and data segments that are relevant to the particular data mining task. Because the raw data may be stored in different formats and in different databases, a large amount of time may be spent on data preprocessing. Postprocessing encompasses all the operations one must perform to make the data mining results more accessible and easier to be interpreted by analysts. For example, spurious results can be filtered according to a variety of measures. Visualization techniques may also be employed to help analysts explore and understand the data mining results.

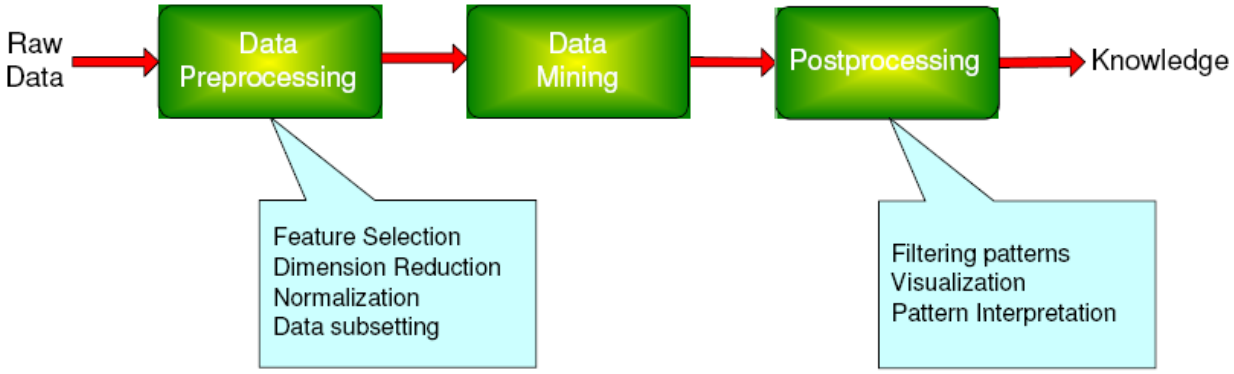


Figure 1. The process of knowledge discovery in databases (KDD).

Data Mining Tasks

In general, data mining tasks can be broadly divided into two major categories:

Predictive Task, which uses certain features of the data to predict the values of other features. Examples include predicting online users who will make a purchase at an e-commerce site, forecasting the future prices of various stocks,and predicting the functionalities of protein structures in human genome.

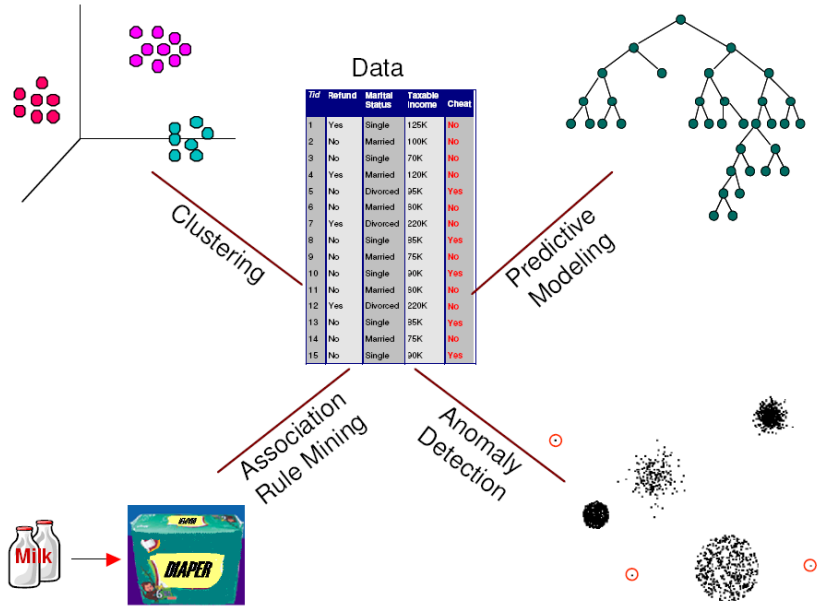
Descriptive Task, which attempts to find human-interpretable patterns that describe certain characteristics of the data. Examples include finding patterns that characterize the different cyber-attacks launched against a computer network, identifying groups of documents that share a common topic, and discovering the climate and ecological factors that influence the weather pattern. A description for each task is given below.

Predictive Modeling. The goal of this task is to build a model for predicting the values of certain features. The input data for a predictive modeling task consists of two types of features (or

variables): (1) Explanatory variables,which correspond to the features used for making the prediction, and (2) target variable, which corresponds to the feature whose value is to be predicted. For example, the explanatory variables for predicting online customers at an ecommerce Web site include the demographic features of the users (such as their age, gender, and salary) along with clickthrough features (such as the list of Web pages browsed by each user and the viewing time for each Web page).

The target variable in this case is a binary feature indicating whether the online user will buy a product at the Web site. The predictive modeling task can be further classified into two subcategories: classification and regression.Classification is used to predict the value of a discrete target variable (e.g.,predicting online users who make a purchase at a Web site). In contrast,regression is used to predict the value of a continuous target variables (e.g.,forecasting the future price of a stock).

Association Analysis. The goal of this task is to produce a set of rules describing the sets of features that are "strongly related" to each other. For example, association analysis can be used to identify products that are often bought together by sufficiently many customers, a task that is commonly known as market basket analysis.



Clustering. The goal of this task is to identify "homogeneous" groupings of data points so that data points that belong together in the same cluster are more similar to each other compared to data points from different clusters. Clustering has many practical applications, including segmenting customers based on their interests, categorizing documents based on their topics, and segmenting land areas according to their vegetation cover.

Anomaly Detection. The goal of this task is to discover anomalies or outliers, i.e.,data points that are significantly different than the rest of the data. Examples of this task include identifying cyber-attacks in computer network systems and predicting fraudulent credit card transactions.

What is not Data Mining?

Besides data mining, there are other ways to extract information from large data repositories. Below, we illustrate some examples of activities that are not considered to be data mining:

• Searching the Internet for Web pages about certain topics. This type of activity is considered more as information retrieval than data mining. Nevertheless,the technology employed by the search engines to determine whether a Web page is relevant to a user's query is often considered as part of data mining.For example, clustering of Web pages may facilitate the retrieval task by efficiently indexing the related pages while classification helps to determine which page is relevant to the input query.

• Looking up the address of a person. Directory lookup is simply a database querying operation and thus does not require intelligent automation. In fact,most database management systems already provide facilities to perform this task. Nevertheless, if additional operations are needed to uncover certain information previously unknown to the user, such operations can be collectively regarded as a data mining task. For instance, finding out the address where John Doe lives is not considered as a data mining task. However, finding out certain names are more prevalent in certain parts of the United States is considered as data mining since the task requires more intelligent techniques to efficiently locate the names and regions.A data mining task has certain features that distinguish it from other information extraction activities. First, the amount of data must be sufficiently large to make it difficult for domain experts to manually discover useful knowledge. For example, computing the average GPA scores for five students is not a data mining task.Second, some aspects of intelligent automation are required to extract the useful information. This is why a simple database querying is not considered as a data mining task. In addition, the task must be complex enough such that it cannot be solved using simple pen-and-paper techniques. For example, counting the number of hits a Web page receives is not data mining but predicting how an online user will behave is a data mining task.

Lesson Six

WIFI: 802.11 WIRELESS LANS

Pervasive in the workplace, the home, educational institutions, cafes, airports, and street corners, wireless LANs are now one of the most important access network technologies in the Internet today. Although many technologies and standards for wireless LANs were developed in the 1990s, one particular class of standards has clearly emerged as the winner: the **IEEE 802.11 wireless LAN**, also known as **WiFi**. In this section, we'll take a close look at 802.11 wireless LANs, examining its frame structure, its medium access protocol, and its internetworking of 802.11 LANs with wired Ethernet LANs.There are several 802.11 standards for wireless LAN technology, including 802.11b, 802.11a, and 802.11 g. Table 1 summarizes the main characteristics of these standards. As of this writing (spring 2009), far more 802.11g devices are now offered by access point and LAN card vendors. A number of dual-mode (802.1 la/g) and tri-mode (802.1 la/b/g) devices are also available.
The three 802.11 standards share many characteristics. They all use the same medium access protocol, CSMA/CA, which we'll discuss shortly. All three use the same frame structure for their link-layer frames as well. All three standards have the ability to reduce their transmission rate in order to reach out over greater distances.And all three standards allow for both "infrastructure mode and "ad hoc mode," as we'll also shortly discuss. However, as shown in Table 1, the three standards have some major differences at the physical layer.

Standard	Frequency Range (United States)	Data Rate
802.11b	2.4-2.485 GHz	up to 11 Mbps
802.11a	5.1-5.8 GHz	up to 54 Mbps
802.1 lg	2.4-2.485 GHz	up to 54 Mbps

Table1 ♦ Summary of IEEE 802.11 standards

The 802.11b wireless LAN has a data rate of 11 Mbps and operates in the unlicensed frequency band of 2.4-2.485 GHz, competing for frequency spectrum with 2.4 GHz phones and microwave ovens. 802.1 la wireless LANs can run at significantly higher bit rates, but do so at higher frequencies. By operating at a higher frequency, 802.11a LANs have a shorter transmission distance for a given power level and suffer more from multipath propagation. 802.1 lg LANs, operating in the same lower-frequency band as 802.11b and being backwards compatible with 802.1 lb (so one can upgrade 802.1lb clients incrementally) yet with the higher speed transmission rates of 802.11a, allows users to have their cake and eat it too.
A new WiFi standard, 802.1 In [IEEE 802.1 In 2009], is in the standardization process. 802.1 In uses multiple-input multiple-output (M1MO) antennas; i.e., two or more antennas on the sending side and two or more antennas on the receiving side that are transmitting/receiving different signals [Diggavi 2004]. Although the standard has yet to be finalized, pre-standard products are available, with early tests showing throughput of over 200 Mbps being achieved in practice [Newman 2008].
One important concern with the current draft standard is the manner in which 802.11 n devices will interact with existing 802.1 la/b/g/devices.

The 802.11 Architecture

Figure 1 illustrates the principal components of the 802.11 wireless LAN architecture. The fundamental building block of the 802.11 architecture is the **basic service set (BSS).** A BSS

contains one or more wireless stations and a central **base station**,known as an **access point (AP)** in 802.11 parlance. Figure 1 shows the AP in each of two BSSs connecting to an interconnection device (such as a switch or router), which in turn leads to the Internet. In a typical home network, there is one AP andone router (typically integrated together as one unit) that connects the BSS to the Internet.

As with Ethernet devices, each 802.11 wireless station has a 6-byte MAC address that is stored in the firmware of the station's adapter (that is, 802.11 network interface card). Each AP also has a MAC address for its wireless interface. As with Ethernet, these MAC addresses are administered by IEEE and are (in theory) globally unique.

Wireless LANs that deploy APs are often referred to as **infrastructure wireless LANs**, with the "infrastructure" being the APs along with the wired Ethernet infrastructure that interconnects the APs and a router. Figure 2 shows that IEEE 802.11 stations can also group themselves together to form an adhoc network—a network with no central control and with no connections to the "outside world." Here, the network is formed "on the fly," by mobile devices that

have found themselves in proximity to each other, that have a need to communicate,and that find no preexisting network infrastructure in their location. An ad hoc network might be formed when people with laptops get together (for example, in aconference room, a train, or a car) and want to exchange data in the absence of a centralized AP. There has been tremendous interest in ad hoc networking, as communicating portable devices continue to proliferate. In this section, though, we'll focus our attention on infrastructure wireless LANs.

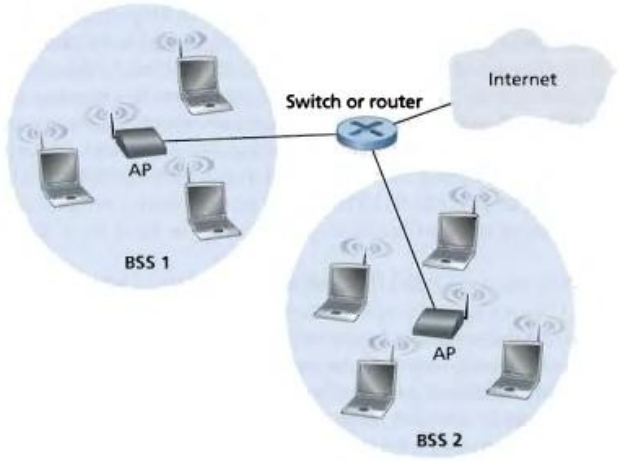


Figure 1 ♦ IEEE 802.11 LAN architecture

Figure 2 ♦ An IEEE 802.11 ad hoc network

Channels and Association

In 802.11, each wireless station needs to associate with an AP before it can send or receive network-layer data. Although all of the 802.11 standards use association,we'll discuss this topic specifically in the context of IEEE 802.1 lb/g.When a network administrator installs an AP, the administrator assigns a one-or two-word **Service Set Identifier (SSID)** to the access point. (When you "view available networks" in Microsoft Windows XP, for example, a list is displayed showing the SSID of each AP in range.) The administrator must also assign a channel number to the AP. To understand channel numbers, recall that 802.11 operates in the frequency range of 2.4 GHz to 2.485 GHz. Within this 85 MHz band, 802.11 defines 11 partially overlapping channels. Any two channels are non-overlapping if and only if they are separated by four or more channels. In particular, the set of channels 1, 6, and 11 is the only set of three non-overlapping channels. This means that an administrator could create a wireless LAN with an aggregate maximum transmission rate of 33 Mbps by installing three 802.1 lb APs at the same physical location, assigning channels I, 6, and 11 to the APs, and interconnecting each of the APs with a switch.

Now that we have a basic understanding of 802.11 channels, let's describe an interesting (and not completely uncommon) situation—that of a WiFi jungle. A **WiFi jungle** is any physical location where a wireless station receives a sufficiently strong signal from two or more APs. For example, in many cafes in New York City,a wireless station can pick up a signal from numerous nearby APs. One of the Aps might be managed by the cafe, while the other APs might be in residential apartments near the cafe. Each of these APs would likely be located in a different IP subnet and would have been independently assigned a channel.

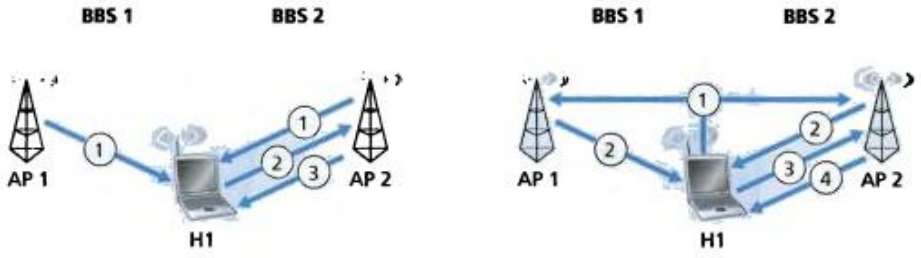
Now suppose you enter such a WiFi jungle with your portable computer,seeking wireless Internet access and a blueberry muffin. Suppose there are five APs in the WiFi jungle. To gain Internet access, your wireless station needs to join exactly one of the subnets and hence needs to **associate** with exactly one of the APs. Associating means the wireless station creates a virtual wire between itself and the AP. Specifically, only the associated AP will send data frames (that is,frames containing data, such as a datagram) to your wireless station, and your wireless station will send data frames into the Internet only through the associated AP. But how does your wireless station associate with a particular AP? And more fundamentally, how does your wireless station know which APs, if any, are out there in the jungle?

The 802.11 standard requires that an AP periodically send **beacon frames**, each of which includes the AP's SSID and MAC address. Your wireless station, knowing that APs are sending out beacon frames, scans the 11 channels, seeking beacon frames from any APs that may be out there (some of which may be transmitting on the same channel—it's a jungle out there!). Having learned about available APsfrom the beacon frames, you (or your wireless host) select one of the APs forassociation.

The 802.11 standard does not specify an algorithm for selecting which of the available APs to associate with; that algorithm is left up to the designers of the 802.11 firmware and software in your wireless host. Typically, the host chooses the AP whose beacon frame is received with the highest signal strength. While a high signal strength is good, signal strength is not the only AP characteristic that will determine the performance a host receives. In particular, it's possible that the selected AP may have a strong signal, but may be overloaded with other affiliated hosts (that will need to share the wireless bandwidth at that AP), while an unloaded AP is not selected due to a slightly weaker signal.

The process of scanning channels and listening for beacon frames is known as **passive scanning** (see Figure 3a). A wireless host can also perform **active scanning**, by broadcasting a probe frame that will be received by all APs within the wireless host's range, as shown in Figure 3b. APs respond

to the probe request frame with a probe response frame. The wireless host can then choose the AP with which to associate from among the responding APs.
After selecting the AP with which to associate, the wireless host sends an association request frame to the AP, and the AP responds with an association response frame. Note that this second request/response handshake is needed with active scanning, since an AP responding to the initial probe request frame doesn't know which of the (possibly many) responding APs the host will choose to associate with, in much the same way that a DHCP client can choose from among multiple DHCP servers .Once associated with an AP, the host will want to join the subnet to which the AP belongs. Thus,the host will typically send a DHCP discovery message into the subnet via the AP in order to obtain an IP address on the subnet. Once the address is obtained, the rest of the world then views that host simply as another host with an IP address in that subnet.
In order to create an association with a particular AP, the wireless station may be required to authenticate itself to the AP 802.11 wireless LANs provide a number of alternatives for authentication and access. One approach, used by many companies, is to permit access to a wireless network based on a station's MAC address. A second approach, used by many Internet cafes, employs usernames and passwords.
In both cases, the AP typically communicates with an authentication server, relaying information between the wireless end-point station and the authentication server using a protocol such as RADIUS [RFC 2865] or DIAMETER [RFC 3588]. Separating the authentication server from the AP allows one authentication server to serve many APs, centralizing the (often sensitive) decisions of authentication and access within the single server, and keeping AP costs and complexity low. We'll see



a. Passive scanning
1. Beacon frames sent from APs
2. Association Request frame sent: H1 to selected AP
3. Association Response frame sent: Selected AP to H1

b. Active scanning
1. Probe Request frame broadcast from H1
2. Probes Response frame sent from APs
3. Association Request frame sent: H1 to selected AP
4. Association Response frame sent: Selected AP to H1

Figure 3 ♦ Active and passive scanning for access points

Lesson Seven

BIOMETRIC TECHNOLOGIES AND VERIFICATION SYSTEMS

Biometric technologies are automated methods for identifying a person or verifying a person's identity based on the person's physiological or behavioral characteristics. Physiological characteristics include fingerprints,hand geometry, and facial, voice, iris, and retinal features; behavioral characteristics include the dynamics of signatures and keystrokes. Biometric technologies capture and process a person's unique characteristics, and then verify that person's identity based on comparison of the record of captured characteristics with a biometric sample presented by the person to be verified. After many years of research and development, biometric technologies have become reliable and cost-effective, and acceptable to users. However, new applications of biometrics are being somewhat successfully implemented in more secure travel documents, visas, and personal identity verification cards. These applications help to safeguard valuable assets and information and contribute to the safety and security of automated transactions, but have fallen short of strengthening homeland security.

Both public and private sectors are looking for reliable, accurate, and practical methods for the automated verification of identity. And they are using biometric technologies in a wide variety of applications, including health and social service programs, passport programs, driver licenses, electronic banking,investing, retail sales, and law enforcement (such as it is). Verification systems are usually characterized by three factors:

- *Something that you know, such as a password;*
- *Something that you have, such as an ID badge;*
- *Something that you are, such as your fingerprints or your face.*

Systems that incorporate all three factors are stronger than those that use only one or two factors. Verification using biometric factors can help to reduce identity theft and the need to remember passwords or to carry documents, which can be counterfeited. When biometric factors are used with one or two other factors, it is possible to achieve new and highly secure identity applications.

For example, a biometric factor can be stored on a physical device, such as a smart card that is used to verify the identification of an individual. Today, the identification cards that are issued to employees for access to buildings and to information, and the cards that are used for financial transactions, often include biometric information.

How Biometric Technologies Work

Biometric technologies vary in complexity, capabilities, and performance, but all share several elements. Biometric identification systems are essentially pattern recognition systems. They use acquisition devices such as cameras and scanning devices to capture images, recordings, or measurements of an individual's characteristics,and use computer hardware and software to extract, encode, store,and compare these characteristics. Because the process is automated, biometric decision making is generally very fast, in most cases taking only a few seconds in real time .

Depending on the application, biometric systems can be used in one of two modes: verification or identification. Verification (also called authentication) is used to verify a person's identity—that is, to authenticate that individuals are who they say they are. Identification is used to establish a person's identity—that is, to determine who a person is. Although biometric technologies measure different characteristics in substantially different ways, all biometric systems involve similar processes that can be divided into two distinct stages: enrollment and verification or identification.

Enrollment

In enrollment a biometric system is trained to identify a specific person. The person first provides an identifier, such as an identity card. The biometric is linked to the identity specified on the identification document. He or she then presents the biometric (fingertips, hand, or iris) to an acquisition device. The distinctive features are located and one or more samples are extracted, encoded, and stored as a reference template for future comparisons. Depending on the technology, the biometric sample may be collected as an image, a recording, or a record of related dynamic measurements.

How biometric systems extract features and encode and store information in the template is based on the system vendor's proprietary algorithms. Template size varies depending on the vendor and the technology. Templates can be stored remotely in a central database or within a biometric reader device itself; their small size also allows for storage on smart cards or tokens .Minute changes in positioning, distance, pressure, environment, and other factors influence the generation of a template, making each template likely to be unique as an individual's biometric data are captured and a new template is generated. Consequently, depending on the biometric system, a person may need to present biometric data several times in order to enroll. The reference template may then represent an amalgam of the captured data, or several enrollment templates may be stored. The quality of the template or templates is critical in the overall success of the biometric application. Because biometric features can change over time, people may have to re-enroll to update their reference template. Some technologies can update the reference template during matching operations .

The enrollment process also depends on the quality of the identifier the enrollee presents. The reference template is linked to the identity specified on the identification document. If the identification document does not specify the individual's true identity, the reference template will be linked to a false identity .

Thus, verification is a one-to-one comparison of the biometric sample with the reference template on file. A reference template is the enrolled and encoded biometric sample of record for a user. Identification makes a one-to-many comparison to determine a user's identity. It checks a biometric sample against all the reference templates on file. If any of the templates on file match the biometric sample, there is a good probability the individual has been identified .

Verification

In verification systems, the step after enrollment is to verify that a person is who he or she claims to be (the person who enrolled). After the individual provides whatever identifier he or she enrolled with, the biometric is presented, and the biometric system captures it, generating a trial template that is based on the vendor's algorithm. The system then compares the trial biometric template with this person's reference template, which was stored in the system duringenrollment, to determine whether the individual's trial and stored templates match (see Figure 1) .

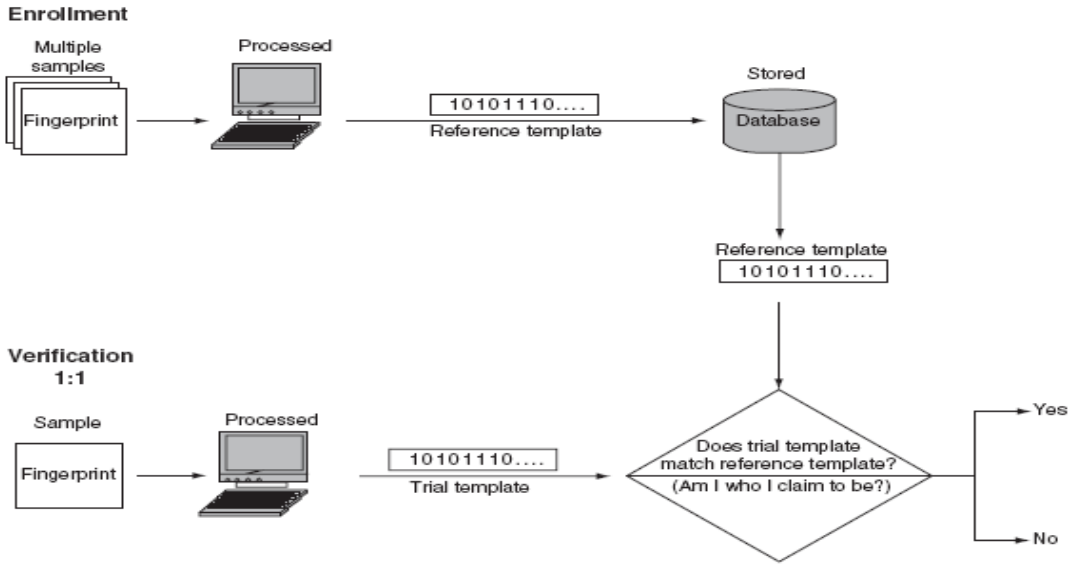


Figure 1- Verification proccess

Verification is often referred to as 1:1 (one-to-one) matching. Verification systems can contain databases ranging from dozens to millions of enrolled templates but are always predicated on matching an individual's presented biometric against his or her reference template. Nearly all verification systems can render a match/no-match decision in less than a second. A system that requires employees to authenticate their claimed identities before granting them access to secure buildings or to computers is a verification application .

Identification

In identification systems, the step after enrollment is to identify who the person is. Unlike verification systems, no identifier need be provided. To find a match,instead of locating and comparing the person's reference template against his or her presented biometric, the trial template is compared against the stored reference templates of all individuals enrolled in the system (see Figure 2) .

Identification systems are referred to as 1:N(one-to-N, or one-to-many) matching because an individual's biometric is compared against multiple biometric templates in the system's database.

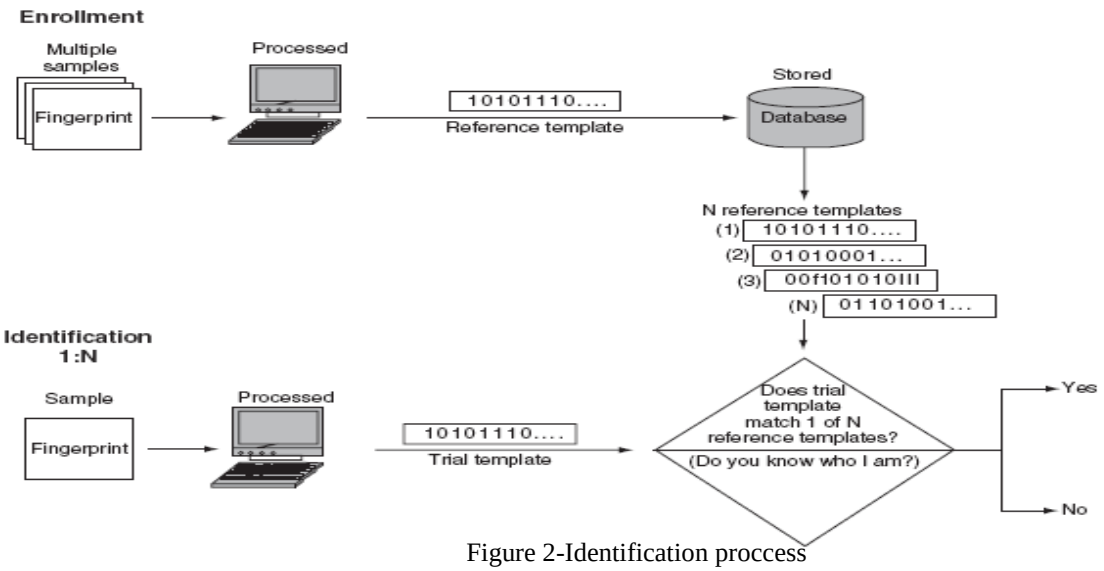


Figure 2-Identification proccess

There are two types of identification systems: positive and negative. Positive identification systems are designed to ensure that an individual's biometric is enrolled in the database. The anticipated result of a search is a match. A typical positive identification system controls access to a secure building or secure computer by checking anyone who seeks access against a database of enrolled employees. The goal is to determine whether a person seeking access can be identified as having been enrolled in the system .

Negative identification systems are designed to ensure that a person's biometric information is not present in a database. The anticipated result of a search is a nonmatch. Comparing a person's biometric information against a database of all who are registered in a public benefits program, for example, can ensure that this person is not "double-dipping" by using fraudulent documentation to register under multiple identities .

Another type of negative identification system is a surveillance system that uses a watch list. Such systems are designed to identify people on the watch list and alert authorities for appropriate action. For all other people, the system is to check that they are not on the watch list and allow them normal passage.The people whose biometrics are in the database in these systems may not have provided them voluntarily. For instance, for a surveillance system, the biometrics may be faces captured frommug shots provided by a law enforcement agency .

References

1 Windows Server 2003 Network Infrastructure / Craig Zacker with Microsoft Corporation.

2 Microsoft Windows XP Networking Inside Out / Curt Simmons.

3 Microsoft C# Professional Projects/ Geetanjali Arora

4 A Survey of Reverse Engineering Tools for the 32-Bit Microsoft Windows Environment/ RAYMOND J. CANZANESE

5 Data mining……

6 Computer Networking A Top Down Aproach/Kurose & Ross

7 Biometric technologies and verification systems / John Vacca.