

12

NETWORK SECURITY

INTRODUCTION

Security for microcomputer LANs has increased in importance, and more vendors are supplying LAN security systems. As end-users become more aware of the value of the vast amounts of data being accumulated and the need to protect that data, more users adopt such systems.

Newspapers carry stories almost every week about some computer network being penetrated, either for financial gain or as a prank. Most of these break-ins involve large corporate networks and wide area networks. As local area networks proliferate and tap into national and international data communication systems, these local networks will also become targets.

Companies that do sensitive work, such as those with defence contracts, are often heavily involved in data security. Other companies may be aware only of the threat, but not of their own vulnerability. Most analysts agree that businesses and institutions, such as schools, will have to suffer a loss through theft or vandalism before they actually establish measures to protect their data.

A computing network, like any other valuable, shared resource, is subject to breaches of security. Such breaches can be accidental or intentional, and their effects on network operations can range from harmless to irritating to devastating.

Security is a critical issue to those planning, managing or using a LAN. It is also a very complex issue. Security is a component of overall network reliability. However, reliability depends largely upon the dependability of network hardware, software and technology. In contrast, the security of a network depends almost exclusively upon the behaviour of that network's authorised users, managers and their guests.

Security, like reliability, is best addressed as part of an overall network strategy. Security concerns must be balanced by other factors that affect the network and its users. Users and managers must therefore discover and implement methods that improve network security without infringing upon users' work patterns or implying that all users are suspected violators of security.

Users have other concerns that network security methods must address as well. Users must be reassured that they can collaborate on projects and share information without being spied upon by managers or other users. Well-implemented password protection schemes can provide much of this reassurance. Managers must also demonstrate to users that procedures for tracking user work patterns on the network are used to improve security and reliability, and not merely to keep a closer eye on users or their activities.

Security methods must be selected with care and implemented with the full cooperation and knowledge of authorized users if security is to be assured (see Figure 12.1). A first step toward these goals is a definition of network security:

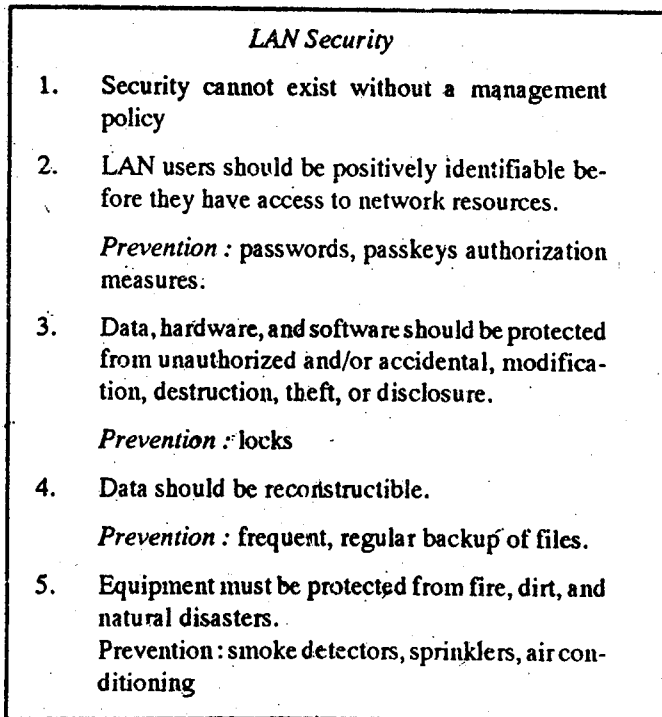


Fig. 12.1 : Essentials of LAN Security

WHAT NETWORK SECURITY MEANS

All the features of electromagnetic media that are desirable to a user also make this media vulnerable to theft and damage. Information stored on disk is easily copied,

altered and erased. As larger amounts of critical data are stored in this way, the significance of the problem grows.

A stand-alone personal computer is easy to secure. You simply put your diskettes in a safe and store your computer in a locked closet. But when you attach that computer to a network of computers, security becomes more complicated. Even a "local" network probably will spread out through several offices, with connecting cables running in ceilings and floors, and in halls and basements. A thief or vandal can tap into any one of a dozen or more spots on the network, many secluded from normal observation. But tapping into the network from some secluded spot on the cable is not usually necessary. A person can simply log-on to a convenient PC and steal or damage data at will. Unfortunately, the easier a system is to use, the easier it is to misuse.

Like any other kind of insurance, data security involves trade-offs. You must weigh the cost of the potential loss against the cost of protection as well as any inconvenience the security measures may cause. The first thing to do in planning your data security programme is to put a value on the data you are going to protect.

In general, a secure network is one that is resistant to disruptions caused by unauthorised network use. Such a network is designed and operated to minimize unauthorised use and can recover from disruptions easily and completely should unauthorised users evade safeguards.

Network security can be defined as the protection of network resources against unauthorised disclosure, modification, utilisation, restriction or destruction. Security has long been an object of concern and study for both data processing systems and communication facilities. With computer networks, these concerns are combined. And for local networks, the problems may be most acute.

Consider a full-capacity local network, with direct terminal access to the network and data files and applications distributed among a variety of processors. The local network may also provide access to and from long-haul communication and be part of an internet. The complexity of the task of providing security in such an environment is clear. The subject is a broad one, and encompasses physical and administrative controls as well as automated ones.

This general definition of a secure network is the foundation upon which you must build a definition that fits your work group's specific requirements and constraints. An effective definition requires careful assessment of needs by you, your colleagues and your managers.

RISK ANALYSIS

Before you can realistically decide how much time and money to invest in data security, you must quantify the risk. Risk analysis has been elevated to a precise discipline. For our purposes, we would not need to examine formulas or other exact methods of quantifying every risk associated with networked data. But we can look briefly at some of the elements of risk analysis. These can help you to develop a preliminary description of your data's value and potential for loss.

First, you will want to determine two values, in rupees, for the information stored in your data system. One is the cost of re-creating the data; the other is the value of lost business if a competitor should gain access to your data.

These two figures should be easy to obtain or at least to estimate. Many smaller companies have never considered the potential loss of their stored data. If nothing else, such an appraisal should encourage the use of data back-up and the insistence on serious password security procedures.

Next, you should identify any possible threats to your data. If your data has little or no monetary value to a competitor, then there is probably little risk of theft. On the other hand, the value of your data to a competitor may be great, with the risk of theft proportionally high.

The physical volume of valuable data is another element to consider. If the volume and diversity of the data are extensive, the chance of a total loss by theft is reduced. A related calculation is the frequency of potential thefts. This figure can be difficult to predict unless you have compiled a history of losses over some period of time. Law enforcement agencies and some trade associations keep extensive records of thefts, defined by type of business, kind of penetration and value of loss. Contacting these groups may turn up sufficient data to allow you to make an intelligent prediction of risk. In addition, you should make a detailed study of any active attacks on your data so that you can estimate the cost of countering a similar attack.

Vandalism is another threat, possibly more serious than theft because the frequency of vandalism is often greater. A discontented employee may decide to "get even" by destroying or altering important files. Or an act of vandalism may be done simply as a prank or game, just to see if it can be done.

After you calculate the value of your data and the types of risks, the final element in risk analysis is the data's vulnerability. Remote access is one factor that causes data to become more easily available and vulnerable. When people can access your network remotely, the potential for loss increases.

On a local basis, the risk to data goes up when the network's contents are generally known. The capability to see those contents (for example, files servers, and other resources) is controlled partially by the operating system and partially by the site administration.

Making a risk analysis will enable you to answer many questions about where risks are greatest and how much money and procedural inconveniences are necessary to thwart these threats. Next, you should consider steps for building a secure data network.

TYPES OF THREATS

A publication of the National Bureau of Standards identified some of the threats that have stimulated the upsurge of interest in security:

1. Organised and intentional attempts to obtain economic or market information

from competitive organisations in the private sector.

2. Organised and intentional attempts to obtain economic information from government agencies.
3. Inadvertent acquisition of economic or market information.
4. Inadvertent acquisition of information about individuals.
5. Intentional fraud through illegal access to computer data banks with emphasis, in decreasing order of importance, on acquisition of funding data, economic data, law enforcement data and data about individuals.
6. Government intrusion on the rights of individuals.
7. Invasion of individual rights by the intelligence community.

These are examples of specific threats that an organisation or an individual (or an organisation on behalf of its employees) may feel the need to counter. The nature of the threat that concerns an organisation will vary greatly from one set of circumstances to another. Fortunately, we can approach the problem from a different angle by looking at the generic types of threats that might be encountered.

Table 12.1 lists the types of threats that might be faced in the context of network security. The threats can be divided into the categories of passive threats and active threats (see Figure 12.2).

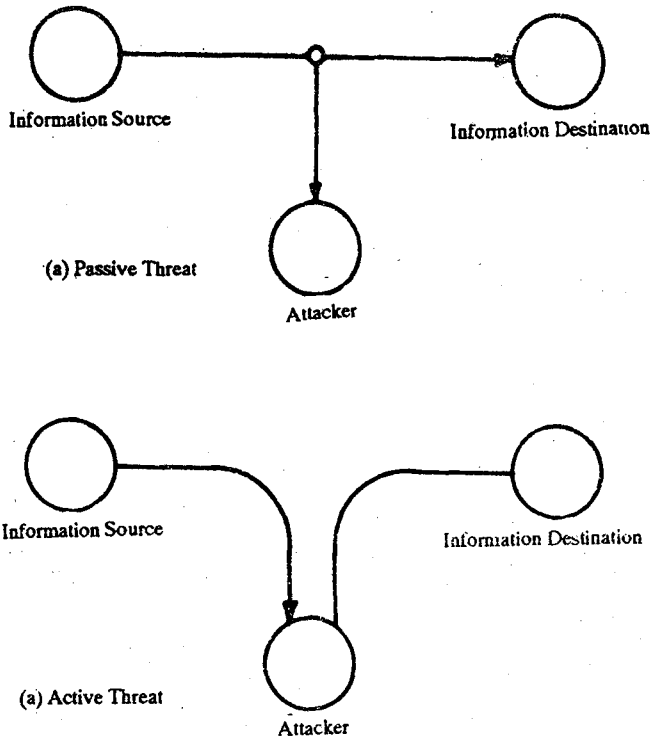


Fig. 12.2 : Passive and Active Communications Security Threats

TABLE 12.1 : Potential Network Security Threats

PASSIVE THREATS

The monitoring and/or recording of data while the data are being transmitted over a communication facility.

RELEASE OF MESSAGE CONTENTS

Attack can read the user data in messages.

TRAFFIC ANALYSIS

The attacker can read packet headers, to determine the location and identity of communicating hosts. The attacker can also observe the length and frequency of messages.

ACTIVE THREATS

The unauthorised use of a device attached to a communication facility to alter transmitting data or control signals or to generate spurious data or control signals.

MESSAGE-STREAM MODIFICATION

The attacker can selectively modify, delete, delay, reorder and duplicate real messages.

The attacker can also insert counterfeit messages.

DENIAL OF MESSAGE SERVICE

The attacker can destroy or delay most or all messages.

MASQUERADE

The attacker can pose as a real host or switch and communicate with another host or switch to acquire data or services.

1. Passive Threats

These are in the nature of eavesdropping or monitoring of the transmissions of an organisation. The goal of the attacker is to obtain information that is being transmitted. Two types of threats are involved here: release of message contents and traffic analysis.

The threat of release of message contents is clearly understood by most managers. A telephone conversation, an electronic mail message or a transferred file may contain sensitive or confidential information. We would like to prevent the attacker from learning the contents of these transmissions.

The second passive threat, traffic analysis, is more subtle and often less applicable. Suppose that we had a way of masking the contents of messages or other information traffic so that an attacker, even if he or she captured the message, would be unable to extract the information from the message. The common technique for doing this is encryption, discussed at length subsequently. If we had such protection in place, it might still be possible for an attacker to observe the pattern of these messages. The attacker can determine the location and identity of communicating hosts and can

also observe the frequency and length of messages being exchanged. This information might be useful in guessing the nature of the communication that is taking place.

Passive threats are very difficult to detect since they do not involve any alteration of the data. However, it is feasible to prevent these attacks from being successful. Thus the emphasis in dealing with passive threats is on prevention and not detection.

2. Active Threats

The second major category of threat is active threats. These involve some modification of the data stream or the creation of a false stream. We can subdivide these threats into three categories: message-stream modification, denial of message service and masquerade.

Message-stream modification simply means that some portion of a legitimate message is altered, or that messages are delayed, replayed or reordered, in order to produce an unauthorised effect. For example, a message meaning "Allow J. N. Saxena to read confidential file accounts" is modified to mean "Allow F. C. Bansal to read confidential file accounts."

The **denial of service** prevents or inhibits the normal use or management of communication facilities. This attack may have a specific target; for example, an entity may suppress all messages directed to a particular destination (e.g. the security audit service). Another form of service denial is the disruption of an entire network, either by disabling the network or by overloading it with messages so as to degrade performance.

A **masquerade** takes place when one entity pretends to be a different entity. A masquerade attack usually includes one of the other two forms of active attack. Such an attack can take place, for example, by capturing and replaying an authentication sequence.

Active threats present the opposite characteristics of passive threats. Whereas passive attacks are difficult to detect, measures are available to prevent their success. On the other hand, it is quite difficult to absolutely prevent active attacks, since this would require physical protection of all communication facilities and paths at all times. Instead, the goal with respect to active attacks is to detect these attacks and to recover from any disruption or delays caused by the attack. Because the detection has a deterrent effect, this may also contribute to prevention.

DETERMINING WHAT SECURE MEANS TO YOU AND YOUR LAN

To arrive at a specific definition of security for your LAN, you and your colleagues must first examine your current network or network plans to identify points of vulnerability. Where points of vulnerability occur depends greatly on the work and network use patterns of every member of your work group. An initial challenge, thus, is to determine these patterns accurately without interfering with them.

If you already have a network, your group will have to decide whether written surveys, personal interviews, software that tracks network access by user, or some other method is best for gathering this information. If you are still in the planning stages, you and your group will have to gather the same information about each independent personal computer (PC) user and use the data to hypothesise points of network vulnerability. A consultant may be helpful with this step.

Every network environment is different, with a different list of specific points of vulnerability. However, most environments have certain vulnerable points. Be sure not to overlook these areas in determining your own environment's particular potential weaknesses.

SECURING WORKSTATIONS AND SERVERS

Like LANs themselves, strategies that address security begin on users' desktops, with their workstations. To protect against both accidental and intentional breaches of network security, users must develop good workstation-protection habits.

One simple habit is turning off workstations when leaving for the evening or weekend, so the screens do not attract wandering eyes and hands. Keeping boot (start-up) disks in a non-obvious drawer instead of on a desk or in the workstation's floppy drive also reduces the likelihood of unauthorised access via an authorised user's workstation.

Physical locks are also available for disk drive doors, keyboards and workstations or PC system units. Some of these locks impede both access and theft. LAN users in large or open-office environments should be encouraged to use these additional security measures and not to defeat them by keeping the keys in their unlocked desks.

It is important to note that in many organisations, the most serious threat to workstation security is not unauthorised users with malicious intent. A larger problem is unauthorised access to user workstations by guests or children of authorised users. These legitimate users often sit their charges in front of an absent user's workstation, to play or explore while the worker works.

This problem is most acute during off hours, when network supervision is minimal or absent. Some companies report a similar problem with after-hours office cleaning staff bringing in and playing unauthorised games on PCs connected to a network. Practices like these must be detected and discouraged to prevent serious network problems caused by well-intended but untrained people.

Servers represent another point of potential vulnerability, especially if they are non-dedicated and also used as workstations. A single-user problem on a combined workstation-server can become a network-wide problem. In addition, even a dedicated server can be mistaken for a workstation if it has a keyboard, floppy disk drive and a screen attached.

The more critical your network is to your business, the more seriously you and your colleagues must work to secure your servers. Removal of the keyboard from each

PC-based server is a good first step. You may also want to put warning signs on servers or to secure them behind locked doors, depending upon their configuration and susceptibility to unauthorised access.

SECURING NETWORK PASSWORDS

Another point of vulnerability under direct user supervision is the passwords that allow access to the network itself, as well as to specific resources, such as particular servers, programmes, or files. Users remember their passwords better when they choose their own, so assignment of random passwords is to be avoided in most situations. However, users must be encouraged to use a bit of creativity when selecting their passwords to make them difficult for unauthorised users to guess or discover accidentally.

You and your colleagues should choose as passwords random numbers or word combinations that are not obvious, but have enough personal significance to be remembered easily. Such a password is less likely to be guessed or discovered and is a more effective security measure than a password based on your telephone number, your birthday or a loved one's name.

You and your colleagues must also implement routines for changing your passwords regularly. Some network managers automatically invalidate any passwords more than 30 days old, forcing users to select new ones at least once a month. Your network's security and reliability could be enhanced simultaneously if you and your colleagues changed your personal passwords each time you made complete back-up copies of your network files.

Needless to say, some users write down their passwords or store them in some electronic note file. If these users leave the notes where others can find them, all the security you and your colleagues are trying to implement can be rendered useless. Encourage your co-workers to treat their network passwords like credit card numbers or access codes for automated teller machines and to protect them with at least as much vigilance.

SECURING FILES AND PROGRAMMES

Users can also help protect against unauthorised access to network files and programmes. Keep master and boot copies of programmes on write-protected disks and, if possible, use passwords to protect your work group's network or application software. When copies of important files are stored on easily removable media such as floppy disks or tape cartridges, restrict access to these media by using locks and keys, sign-in and sign-out lists, supervisor monitoring or other measures. These practices reduce the possibility of accidental or malicious erasure or modification of important files.

Files must also be protected while they are in use on a network. Users must strive always to open and close files according to the procedures required by their network and application software. Otherwise, network file directories can become

incorrect or corrupted, and larger problems can result. Most network software offers some protection against these problems, but good user habits are the best safeguards.

Some network programmes require the insertion of key disks into workstation floppy drives to qualify legitimate users for access to programmes and files. Where these disks are in use, they must be protected and not widely distributed or duplicated. The use of third-party programmes that eliminate the need for key disks must also be weighed against the increased security risk that these disks can represent.

Networks must also be protected from unauthorised programmes, such as game programmes or other personal software. Unauthorised programmes can contaminate your network with annoying or highly destructive software viruses.

You and your colleagues should avoid bringing unauthorised software into contact with your network. Whether a harmless game or your own copy of a programme your work group uses, any software not supplied through your network's usual channels should be viewed as a potential source of harm to your network.

LEVELS OF SECURITY

There is no such thing as 100% security. With enough skill and enough time to complete the job, a perpetrator can defeat any security measure.

Of the two security elements of skill and time, the most dependable protection is time. If you can make certain that a break-in will be a time-consuming project for a thief, you have gone a long way in protecting your data. Therefore, all serious security systems are layered with not one but several security measures (see Figure 12.3). For a local area network the following strategies should be considered:

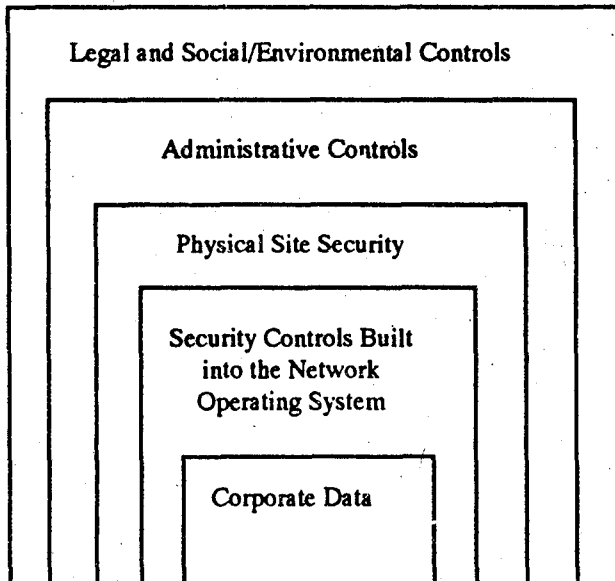


Fig. 12.3 : Layers of Data Security

1. Physical security
2. Access Control
3. Personal identification
4. Encryption
5. The diskless PC
6. Protection against cable radiation
7. Call-back security

1. PHYSICAL SECURITY

Data security can take many forms. The simplest is physical security, which may be a lock on the computer or a guard at the door. With physical security, a would-be thief must attack and defeat your security measures before becoming a threat to the data.

Locks can set up barriers from the back door to the office door to the computer itself. Key locks are now provided for IBM PC/ATs and compatibles. The lock interrupts the power to the display and keyboard, while still allowing the terminal to remain on-line. Turning the key powers up the user interfaces; the key cannot be removed while the system is on. This kind of physical security is available for personal computers.

An alarm system works in partnership with your physical security measures. Locking devices are designed to increase the time needed for penetration. Alarms put an effective limit on the amount of time available. Professional criminals do not run when they hear an alarm or when they think they have tripped a silent alarm. Most know precisely how much time they have before the police arrive. If they cannot get through the security system's physical barriers in the time available, then the criminals will abandon the effort.

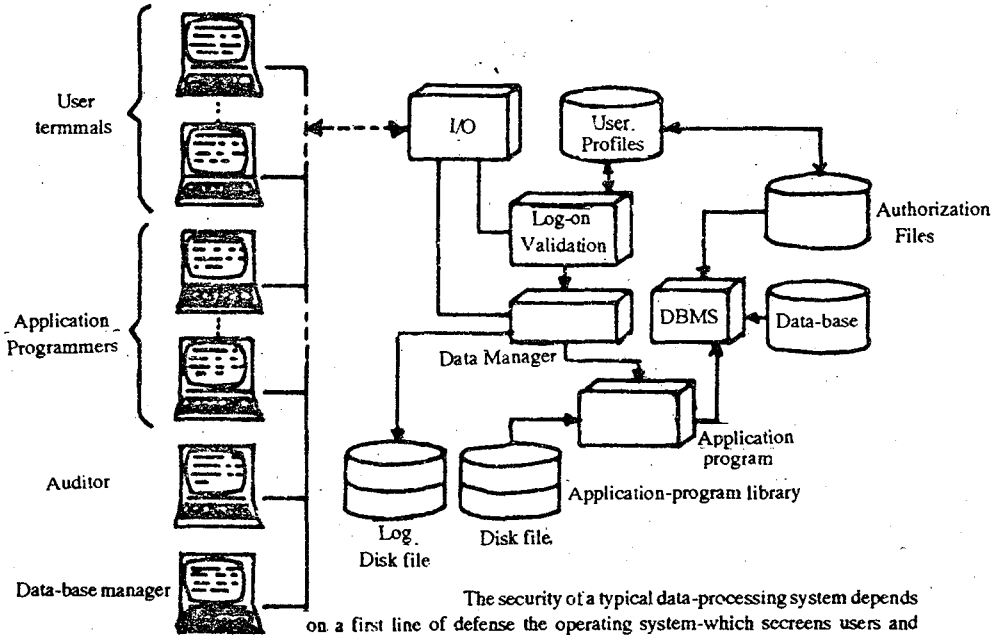
2. ACCESS CONTROLS

The purpose of access controls is to ensure that only authorised users have access to the system and its individual resources and that access to and modification of particular portions of data is limited to authorised individuals and programmes.

Figure 12.4 depicts, generically, the measures taken to control access in a data processing system. They fall into two categories: first those associated with the user or group of users and, second, those associated with the data. In what follows, we elaborate on these concepts and extend them to the local networking environment.

The control of access by user is referred to as authentication. A quite common example of this on a time-sharing system is the user log-on, which requires both a user ID and a password. The system will only allow a user to logon if that user's ID is known to the system and if the user knows the password associated by the system with that ID. This ID/password system is a notoriously unreliable method of

access control. Users can forget their passwords, and accidentally or intentionally reveal their password. Also, the ID/password file is subject to penetration attempts.



The security of a typical data-processing system depends on a first line of defense the operating system-which screens users and transactions and on further screening by the data-base management system to validate queries to the data base.

Fig. 12.4 : Data Processing System Security

No cost-effective method of overcoming this problem exists. Exotic techniques such as voiceprints, fingerprints and hand geometry analysis may be foolproof but are at present prohibitively expensive. Simple measures that can be taken now are to change passwords frequently and to maintain tight multiple measures of security over the ID/password directory. One additional measure that is cost effective is to associate ID's with terminals rather than users and hard wire the code into the terminal. This changes an administrative/software security problem into a physical security problem. However, if it is desirable to allow one-to-many and/or many-to-one relationships between users and terminals, this technique is ineffective.

The problem of authentication is compounded over a multiaccess medium LAN. The log-on dialogue must take place over the communication medium and eavesdropping is a potential threat. One approach to protection would be to certify that each NIU can capture only data addressed to it. This is no easy task. Another approach is to encrypt the ID/password data. User and user group authentication can be either centralised or distributed. In a centralised approach the network provides a log-on service. Distributed authentication treats the network as a transparent communication link, and the usual log-on procedure is carried out by the destination host. Of course, the security concerns for multiaccess media must still be addressed.

In fact, in many local networks, two levels of authentication will probably be used. Individual hosts may be provided with a log-on facility to protect host-specific

resources and applications. In addition, the network as a whole may have protection to restrict network access to authorised users. This two-level facility is desirable currently for the common case, in which the local network connects disparate hosts and simply provides a convenient means of terminal-host access. Future integrated networks (in the OSI sense) may require only a network-level scheme.

Following successful authentication, the user has been granted access to one or a set of hosts and/or processes. This is generally not sufficient for a system that includes sensitive data in its data base. Through the authentication procedure, a user can be identified together with a profile that specifies permissible operations and file accesses. The operating system can enforce rules based on the user profile. The database management system, however, must control access to specific portions of records. For example, it may be permissible for anyone in administration to obtain a list of company personnel, but only selected individuals may have access to salary information. The issue is more than just one of level of detail. Whereas the operating system may grant a user permission to access a file or use an application, following which there are no further security checks, the database management system must take a decision on each individual access attempt. That decision will depend not only on the user's identity but also on the specific parts of the record being accessed, and even on the information already divulged to the user.

A general model of access control as exercised by a data base management system is that of an access matrix (see Table 12.2). One axis of the table consists of identified subjects that may attempt data access. Typically, this list will consist of individual users or user groups, although access could be controlled for terminals, hosts or processes instead of or in addition to users. The other axis lists the objects that may be accessed. At the greatest level of detail, objects may be individual data fields. More aggregate groupings, such as records, record types, or even the entire data base may also be objects in the matrix indicates the access rights of that subject for that object.

TABLE 12.2 : Data Base Access Matrix

| | | Objects | | |
|----------|-------------|---------------|----------------|----------------|
| | | Data Bases... | Record Type... | Record...Field |
| Subjects | Individuals | | | |
| | | | | |
| | | | | |
| | Terminals | | | |
| | | | | |
| | | | | |
| | Hosts | | | |
| | | | | |
| | | | | |
| | Process | | | |
| | | | | |
| | | | | |

Delete, modify,
read, write, execute

In practice, an access matrix is usually sparse, and is implemented by decomposition in one of two ways. The matrix may be decomposed by columns, yielding access control lists. Thus for each object, an access control list lists users and their permitted access opportunities. Decomposition by rows yields capability tickets. A capability ticket specifies authorised objects and operations for a user. Each user has a number of tickets and may be authorised to loan or give them to others. Because tickets may be dispersed around the system, they present a greater security management problem than access control lists.

Network considerations for access control parallel those for authentication. Encryption may be needed to provide secure communications on a LAN. Typically, access control is decentralised, that is, controlled by host-based data management systems. However, if a network data base server exists on a LAN, access control becomes a network service.

3. PERSONAL IDENTIFICATION

A local area network presents some additional security problems because of its dispersed nature and because many people have access to the network. Remote access through modems and telephone lines is used widely on LANs, which makes dispersion essentially infinite. Dispersion thwarts one of the best types of personal-identification security systems.

On most networks the first line of security is personal identification. You physically recognise people who are authorised to be in your office, sitting at a PC. With remote access this kind of identification is impossible. Companies must rely on passwords and classified access schemes to protect their data.

Several techniques can be used to restrict access to authorised users. All these techniques are based on some kind of identification: personal, such as ID badge; key word such as a log-in name and password; or key number.

Badges and personal recognition may not be successful in large companies where everyone is not personally known. In a company with many employees, a counterfeit badge may, in fact, be all that is necessary to penetrate a security system based widely on identification.

PASSWORDS

Password security adds no cost to the network and is potentially a useful security measure. After logging onto the network, the user must type a password. Theoretically, if users must give a password, unauthorised access is prevented. But often the password system is misused and ineffective.

Passwords usually are chosen because they are easily remembered. This, however, also makes them easily guessed. Common assignments include first name for log-in name and last name or title for password. The value of passwords is further diluted when employees give their passwords to others in the organisation. A password often is given out because another employee needs to read a particular file or to perform some task for an absent employee.

Password protection can be improved through both systematised procedures and more sophisticated operating system password utilities. Passwords should be assigned by a network manager, not by the individual. This assignment method reduces the likelihood that someone will identify the password in half a dozen guesses. Many network operating systems have a password utility that allows authorised users to change their own passwords. Such a utility should be deleted from all users' directories and given only to the network supervisor.

Over time, passwords will become generally known, particularly within a small office or department. This decaying security can be stopped by periodically issuing new passwords, say, on a monthly basis. One additional advantage of changing passwords regularly is that employees will take more seriously both the password system and the subject of security.

SECURITY IN LOG-IN

The network operating system should be designed to thwart attempts to break into the system. For one thing, the password should not be "echoed" back to the screen when the user types it in during log-in. The number of times that a password can be attempted should be limited to no more than three tries. After that, the log-in name should be invalidated temporarily, and the network supervisor notified of a failed log-in. An audit trail can also be provided to record the number of password attempts from a given user or station. The presence of the audit trail utility that monitors the password system is a deterrent in itself, especially to malicious or casual vandals.

A sophisticated thief, however, can collect log-in routines and passwords as they are entered often simply by tapping into the network. The network operating system can be enhanced to make this activity more difficult for the thief. Passwords can be encrypted at the workstation and decrypted at the central processor so that the data on the cable is unusable through a tap.

As part of security planning and implementation, an independent analyst should evaluate the security measures, even to the extent of attempting to steal or corrupt a prearranged target file.

4. ENCRYPTION

Earlier, we referred to one of the major security risks on LANs, which uses a multiaccess medium - the risk of eavesdropping. Eavesdropping can be accomplished by programming the NIU to accept packets other than those addressed to it or by physically tapping into the medium. One countermeasure that, properly used, is very effective is to encrypt the data in each packet (i.e. send the data in code).

Encryption is the process of changing intelligible data into unintelligible data; decryption reverses the process. For most local area networks, data encryption is used only when the security threat is substantial.

Ensuring that data is secure in a network environment is more difficult than ensuring the security of physical documents. Typically, data in a network is held in

a common storage facility, and anyone authorised to use the central storage has the potential to access classified files. The best solution to this potential problem is to store the data in an encrypted form. Then, any unauthorised person accessing the file would not be able to read its contents.

Encryption techniques cover a broad range, from simple encryption that guards against accidental disclosure to sophisticated methods which protect against all but the highly trained criminal with an in-depth knowledge of cryptanalysis and considerable deciphering equipment.

Most encryption schemes are based on mathematical operations that are "computationally infeasible". That is, they are based on prime numbers which are so large that even the computational power of a mainframe computer cannot break the code within a practical time period.

Two primary types of encryption exist: link and end-to-end. Link encryption is used to make data unreadable while it is on a point-to-point link, such as between two PCs. Link encryption prevents the casual reading of data.

End-to-end encryption protects data anywhere on the system. This type of encryption corresponds to Layer 4 (the Transport Layer) in the OSI Model. Because Layer 4 is end-to-end, encryption here can provide protection to any number of communication links or intermediate networks.

ENCRYPTION KEYS

Encryption key systems are commonly found on dial-up networks but are also available on LANs. A key is essentially a formula for coding and decoding a message. Keys are carefully distributed to authorised users. In fact, the security of the distribution channel for keys often establishes the security level of a system.

Such a system of secret keys is very difficult and expensive to maintain, especially as the number of participants increases. To overcome these disadvantages, a new key called a "public key" was devised. Public keys may be published openly, and they permit virtually any individual to use a personal public key to code a message and send it to another person. To decode the message, however, the receiver must use a secret key. A secret key consists of two prime numbers that are not published.

One other application of public keys is to authenticate messages. You can use your secret key to encrypt a message and send it to a second person. That person will take your public key and use it to decode the message. If your public key does decode an encoded message, presumably sent by you, then proof has been provided that you did indeed send the message. In other words, the public key is an electronic signature.

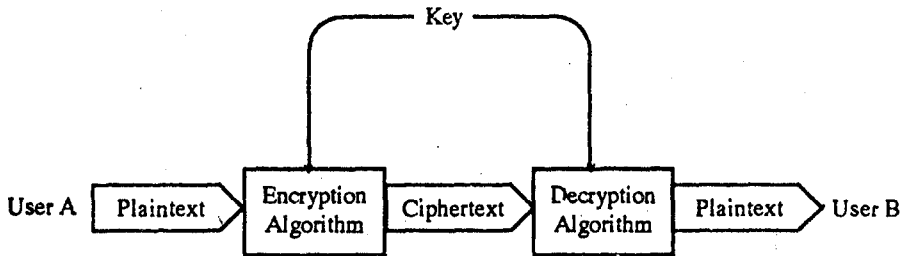
All keys are factorable and, therefore, limited in their level of security. Over the last few years, a debate has been going on about how complex a key should be. Generally, any encryption system provides file privacy against casual perusal. But encryption systems can go far beyond providing file privacy. The Data Encryption Standard (DES) is an encryption system designed by IBM and adopted by the National

Bureau of Standards in 1977. Using an encryption system that conforms to the DES standard generally is considered sufficient protection against unauthorised access. That is, most criminals and vandals would not be able to break into a communication system and steal or alter data which has been encrypted according to the DES standard. With the largest and fastest computers available today, however, DES encryption schemes probably are breakable.

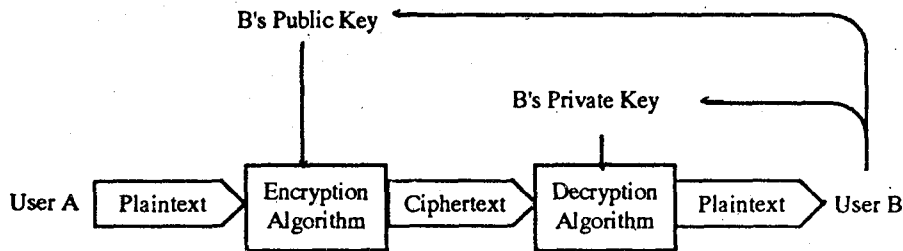
A number of schemes for encryption have been proposed. In this section, we describe two techniques that are good candidates for local network use.

A. CONVENTIONAL ENCRYPTION

Figure 12.5a illustrates the conventional encryption process. The original intelligible message, referred to as plaintext, is converted into apparently random nonsense, referred to as ciphertext. The encryption process consists of an algorithm and a key. The key is a relatively short bit string that controls the algorithm. The algorithm will produce a different output depending on the specific key being used at the time. Changing the key radically changes the output of the algorithm.



(a) Conventional Encryption



(b) Public-Key Encryption

Fig. 12.5 : Encryption

Once the ciphertext is produced, it is transmitted. Upon reception, the ciphertext can be transformed back to the original plaintext by using a decryption algorithm and the same key that was used for encryption.

The security of conventional encryption depends on several factors. First, the encryption algorithm must be powerful enough to make it impractical to decrypt a message on the basis of the ciphertext alone. Beyond that, the security of conventional encryption depends on the secrecy of the key, not the secrecy of the algorithm. That is, it is assumed that it is impractical to decrypt a message on the basis of the ciphertext plus knowledge of the encryption/decryption algorithm. In other words, we do not need to keep the algorithm secret; we only need to keep the key secret.

This feature of conventional encryption is what makes it feasible for widespread use. The fact that the algorithm need not be kept secret means that manufacturers can, and have, developed low-cost chip implementations of data encryption algorithm. These chips are widely available and incorporated into a number of products. With the use of conventional encryption, the principal security problem is maintaining the secrecy of the key. This issue is addressed here.

THE DATA ENCRYPTION STANDARD

The most widely used encryption scheme is based on the data encryption standard (DES), adopted in 1977 by the National Bureau of Standards. For DES, data are encrypted in 64-bit blocks using a 56-bit key. Using the key, the 64-bit input is transformed in a series of steps involving transposition and exclusive-or operations. The result is a 64-bit output in which each bit of output is a function of each bit of the input and each bit of the key. At the receiver, the plaintext is recovered by using the same key and reversing the steps.

The DES has enjoyed widespread use. Unfortunately, it has also been the subject of much controversy as to how secure the DES is. The main concern is in the length of the key, which some observers consider to be too short. To appreciate the nature of the controversy, let us quickly review the history of the DES.

The DES is the result of a request for proposals for a national cipher standard released by the NBS in 1973. At that time, IBM was in the final stages of a project called Lucifer to develop its own encryption capability. IBM proposed the Lucifer scheme, which was by far the best system submitted. It was, in fact, so good that it considerably upset some people at the National Security Agency (NSA), which until that moment had considered itself comfortably ahead of the rest of the world in the still arcane art of cryptography. DES, as eventually adopted, was essentially the same as Lucifer, with one crucial difference: Lucifer's key size was originally 128 bits, whereas the final standard uses a key of 56 bits. What is the significance of the 72 dropped bits?

There are basically two ways to break a cipher. One way is to exploit properties of whatever mathematical functions form the basis of the encryption algorithm to make a "cryptoanalytic" attack on it. It is generally assumed that DES is immune to such attacks, although the role of NSA in shaping the final DES standard leaves lingering doubts. The other way is a brute force attack in which you try all possible keys in an "exhaustive search". That is, you attempt to decrypt ciphertext with every possible 56-bit key until something intelligible pops out. With only 56 bits in the DES key,

there are 2×56 different keys - a number that is uncomfortably small, and becoming smaller as computers get faster.

Whatever the merits of the case, DES has flourished in recent years and is widely used, especially in financial applications. Except in areas of extreme sensitivity, the use of DES in commercial applications should not be a cause for concern by the responsible managers.

COMMERCIAL COMMUNICATIONS SECURITY ENDORSEMENT PROGRAMME

Although DES still has a reasonably useful life ahead of it, it is likely that non-government organisations will begin to look for replacements for what is seen as an increasingly vulnerable algorithm. The most likely replacement is a family of algorithms developed under the NSA commercial COMSEC (communications security) Endorsement Program (CCEP). CCEP is a joint NSA-industry effort to produce a new generation of encryption devices that are more secure than DES, that are low-cost, and that are capable of operating at high data rates. Features of the new CCEP algorithms:

1. The CCEP algorithms are developed by NSA and are classified. Thus the algorithms themselves remain secret and are subject to change from time to time.
2. Industry participants will produce chip implementation of the algorithms, but the NSA maintains control over the design, fabrication and dissemination of chips.

Two types of algorithms come under the CCEP heading. Type I algorithms are designed to protect classified government information. Equipment using Type I CCEP will be available only to government agencies and their designated contractors. Type II algorithms are designed to protect sensitive but unclassified information. Type II gear is intended to replace DES gear. Unlike the Type I modules, which will handle classified information, the Type II equipment is controlled only to the point of sale. Presumably, after a Type II module is built into a computer or communication device and sold by a vendor, the customer can do with it as he or she pleases - short of exporting it overseas.

Although the purpose of developing the Type II equipment, as with the Type I equipment, was to provide a means of protecting government information, the Type II modules are available for use in non-government, private sector applications. As this equipment becomes more widely available, it is likely to become more widely used, at the expense of DES.

KEY DISTRIBUTION

For conventional encryption to work, the two parties to an exchange must have the same key, and that key must be protected from access by others. Furthermore, frequent key changes are usually desirable to limit the amount of data compromised if an attacker learns the key. Therefore, the strength of any cryptographic system rests with the key distribution technique, a term that refers to the means of delivering a

key to two parties that wish to exchange data, without allowing others to see the key. Key distribution can be achieved in a number of ways. For two parties A and B:

1. A key could be selected by A and physically delivered to B.
2. A third party could select the key and physically deliver it to A and B.
3. If A and B have previously and recently used a key, one party could transmit the new key to the other, encrypted using the old key.
4. If A and B each have an encrypted connection to a third party C, C could deliver a key on the encrypted links to A and B.

Options 1 and 2 call for manual delivery of a key, which is awkward. In a distributed system, any given host or terminal may need to engage in exchanges with many other hosts and terminals over time. Thus, each device needs a number of keys, supplied dynamically. The difficulty with Option 3 is that if an attacker ever succeeds in gaining access to one key, then all subsequent keys are revealed.

Option 4 is the most attractive and could be handled from a host facility or network control centre. Figure 12.6 illustrates a possible implementation. For this scheme, two kinds of keys are identified.

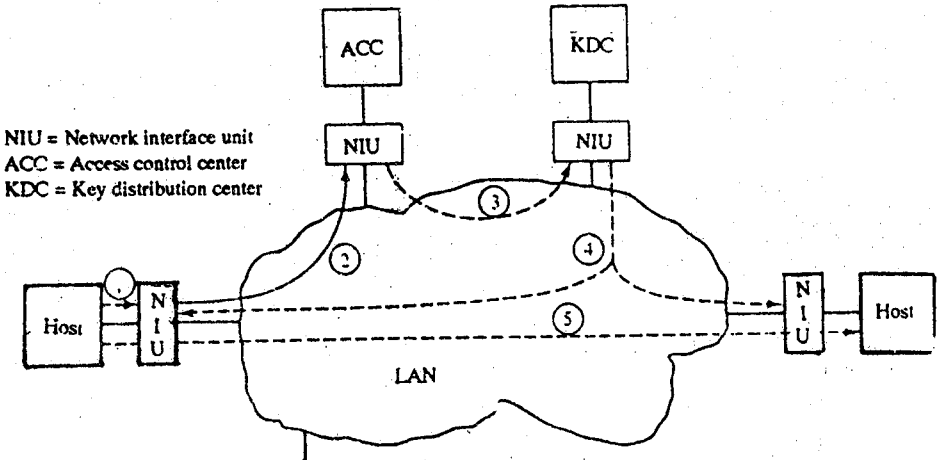
- **Session Key:** When two end-systems (hosts, terminals, etc) wish to communicate, they establish a logical connection (e.g. LLC connection or transport connection). For the duration of that logical connection, all user data are encrypted with a one-time session key. At the conclusion of the session or connection, the session key is destroyed.
- **Permanent Key:** A permanent key is one used between entities for the purpose of distributing session keys.

The configuration consists of the following elements:

- **Access control centre:** The access control centre determines which systems are allowed to communicate with each other.
- **Key distribution centre:** When permission is granted by the access control centre for two systems to establish a connection, the key distribution centre provides a one-time session key for that connection.
- **Network interface unit:** The NIU performs end-to-end encryption and obtains session keys on behalf of its host or terminal.

The steps involved in establishing a connection are shown in Figure 12.6. When one host wishes to set up a connection to another host, it transmits a connection-request packet (1). The NIU saves that packet and applies to the access control centre for permission to establish the connection (2). The communication between the NIU and the access control centre is encrypted using a permanent key shared only by the access control centre and the NIU. The access control centre has one such unique key for each NIU and for the key distribution centre. If the access control centre approves

the connection request, it sends a message to the key distribution centre, asking for a session key to be generated (3). The key distribution centre generates the session key and delivers it to the two appropriate NIUs, using a unique permanent key for each NIU (4). The requesting NIU can now release the connection request packet, and a connection is set up between the two end systems (5). All user data exchanged between the two end systems are encrypted by their respective NIUs using the one-time session key.



NIU = Network interface unit
 ACC = Access control center
 KDC = Key distribution center

1. Host sends packet requesting connection.
2. NIU buffers packet, asks ACC for session key.
3. ACC approves request, commands KDC.
4. KDC distributes session key to both NIUs.
5. Buffered packet transmitted.

Figure 12.6. Key Distribution Across a LAN

Several variations on this scheme are possible. The functions of access control and key distribution could be combined into a single-system. The separation makes the two functions clear and may provide a slightly enhanced level of security. If we wish to let any two devices communicate at will, then the access control function is not needed at all. When two devices wish to establish a connection, one of them applies to the key distribution centre for a session key.

The automated key distribution approach provides the flexibility and dynamic characteristics needed to allow a number of terminal users to access a number of hosts and for the hosts to exchange data with each other. A number of LAN vendors offer some version of the scheme shown in Figure 12.6. It is a powerful and reasonably inexpensive means of enhancing network security.

B. PUBLIC KEY ENCRYPTION

As we have seen, one of the major difficulties with conventional encryption schemes is the need to distribute the keys in a secure manner. A clever way around this requirement is an encryption scheme that, surprisingly, does not require key distribution. This scheme, known as public-key encryption and first proposed in 1976, is illustrated in Figure 12.5b.

For conventional encryption schemes, the keys used for encryption and decryption are the same. This is not a necessary condition. Instead, it is possible to develop an algorithm that uses one key for encryption and a companion but different key for decryption. Furthermore, it is possible to develop algorithms such that knowledge of the encryption algorithm plus the encryption key is not sufficient to determine the decryption key. Thus the following technique will work.

1. Each end-system in a network generates a pair of keys to be used for encryption and decryption of messages that it will receive.
2. Each system publishes its encryption key by placing it in a public register or file. This is the public key. The companion key is kept private.
3. If A wishes to send a message to B, it encrypts the message using B's public key.
4. When B receives the message, it decrypts it using B's private key. No other recipient can decrypt the message since only B knows B's private key.

As you can see, public-key encryption solves the key distribution problem, since there are no keys to distribute! All participants have access to public keys, and private keys are generated locally by each participant and therefore need never be distributed. As long as a system controls its private key, its incoming communication is secure. At any time, a system can change its private key and publish the companion public key to replace its old public key.

A further refinement is needed. Since anyone can transmit a message to A using A's public key, a means is needed to prevent impostors. To develop this scheme, you need to know that public key encryption algorithms are such that the two keys can be used in either order: That is, one can encrypt with the public key and decrypt with the matching private key, or encrypt with the private key and decrypt with the matching public key. Now consider the following scenario: B prepares a message and encrypts it with its own private key, and then encrypts the result with A's public key. On the other end, A first uses its private key and then uses B's public key in a double decryption. Since the message was encrypted with B's private key, it could only come from B. Since it was also encrypted with A's public key, it can only be read by A. With this technique, any two stations can at any time set up a secure connection without a prior secret distribution of keys.

A main disadvantage of public-key encryption compared to conventional encryption is that algorithms for the former are much more complex. Thus, for comparable size and cost of hardware, the public-key scheme will provide much lower throughput. One possible application of public-key encryption is to use it for the permanent key portion of Figure 12.6, with conventional keys used for sessions keys. Since there are few control messages relative to the amount of user data traffic, the reduced throughput should not be a handicap.

Table 12.3 summarises some of the important aspects of conventional and public-key encryption.

TABLE 12.3 : Conventional and Public-Key Encryption

| Conventional Encryption | Public-key Encryption |
|---|--|
| <p>Needed to Work :</p> <ol style="list-style-type: none"> <li data-bbox="47 291 539 388">1. The same algorithm with the same key can be used for encryption and decryption. <li data-bbox="47 414 539 511">2. The sender and receiver must share the algorithm and the key. | <p>Needed to Work :</p> <ol style="list-style-type: none"> <li data-bbox="539 291 1017 414">1. One algorithm is used for encryption and decryption with a pair of keys, one for encryption and one for decryption <li data-bbox="539 414 1017 511">2. The sender and receiver must each have one of the matched pair of keys. |
| <p>Needed for Security :</p> <ol style="list-style-type: none"> <li data-bbox="47 555 539 608">1. The key must be kept secret <li data-bbox="47 635 539 732">2. It must be impossible or at least impractical to decipher a message if no other information is available. <li data-bbox="47 740 539 874">3. Knowledge of the algorithm plus samples of ciphertext must be insufficient to determine the key | <p>Needed for Security :</p> <ol style="list-style-type: none"> <li data-bbox="539 555 1017 643">1. One of the two keys must be kept secret. <li data-bbox="539 643 1017 732">2. It must be impossible or at least impractical to decipher a message if no other information is available. <li data-bbox="539 740 1017 874">3. Knowledge of the algorithm plus of the keys plus samples of ciphertext must be insufficient to determine the key. |

ON-LINE CODERS

The easiest measure to take for LAN security is to attach an encryption device at either end of a communication link. Several companies make such devices and will modify them for specific applications. After the devices are installed, the system is fully transparent to the user. With each person using an encryption box, the message sent between parties will be encrypted while it is on the line.

Another way to set up a system is to place an encryption box between each PC and the network. Then all the data that goes out on the network and all data stored on the hard disk will be encrypted. Ideally, the device can be modified and tuned to provide the speed and security needed. If necessary, a public key system can also be built in.

To show how such a security system might work, let us suppose that we have three groups on a network: administration (admin), accounting and sales. All the data on the network can be encrypted. The administrator can read everything, but accounting and sales can read only their respective files. Each user encrypts the data on an optional basis. With each transmission the encryption device will ask the user whether to transmit in the clear or with encryption. The administrator's device will also ask of the administrator, "Which key do you want: admin, accounting or sales?"

5. THE DISKLESS PC

The power of the PC itself is a potential security threat that should be considered. One of the advantages that the personal computer has over dumb terminals is its local storage capability. Information can be locally manipulated and stored on a PC's floppy diskettes, then transferred to the central storage. From central storage the information can be made available to other users and maintained and backed up properly.

With local storage devices, users can maintain their own back-up system, independent of the central system. The degree of autonomy associated with a personal set of data diskettes is appealing to many users. At the same time, such autonomy creates two threats to data security.

One threat is unintentional. Because two copies of data exist, one on the central disk and one locally, the copies may be updated independently. Eventually, unique data on one version may be lost when the two "copies" are merged.

The other threat is that a local disk drive permits data theft. A person with access to the network and with a local disk drive can copy large amounts of data onto floppy disks in just minutes. The data, then, can be easily hidden and removed from even reasonably secure buildings.

Most network vendors now provide the capability of booting a local PC workstation from a central server so that diskless PCs can be used on the network. Such machines require full-time networking and permit no local storage. A common reason for using diskless PCs is cost. Because diskless PCs require no local floppy controller or disk drive, the cost of a workstation is reduced. But equally important is the increased security offered by a diskless PC.

Take away the disk drive and you take away the means for stealing the data. But you also reduce the power of the PC. In many instances local storage is desirable so that the PC can be used as a stand-alone workstation in the event of a network failure. One answer is to exchange a local floppy for a local hard disk. Then not only would the user have all the benefits of local storage, but local speed and efficiency would improve also. No ready way, however, would be available to copy or remove data.

Diskless PCs have been hampered by software problems. Many application programmes are designed to run only from a local floppy disk drive. Diagnostics and the operating system itself have usually required at least one local drive.

Increasingly, however, software vendors are providing some mechanism for their application packages to be stored on a hard disk and used in a multiuser environment. A company can then make its own decisions about how to configure PCs. Probably the answer will be a variety of configurations to fit particular circumstances.

6. PROTECTION AGAINST CABLE RADIATION

Any time information is transmitted, even through cable, that information can

potentially be intercepted by unauthorised persons. The possibility also exists that a vandal can tamper with data or destroy data files.

Several methods may be used for protecting data while it is on the cable. The first thing to do is to put the cable out of sight. This step should be taken anyway, to prevent damage to the cable and to meet building codes. Security is a secondary benefit. Install cables in protective raceways in areas where penetration is less likely.

A radio signal that is broadcast onto the air waves can easily be intercepted and the information stolen. Such emissions, however, are not limited to broadcasted radio signals. A data cable also radiates intelligible signals, just as a transmitting antenna does. Simple intercept equipment located near the cable can pick up and record these transmissions. More sophisticated devices can intercept the signals a considerable distance from the cable.

The likelihood of signal interception can be eliminated by using a shielded cable, which is a cylinder of braided copper wire that encases the intelligence-carrying wires. If one shield does not reduce emissions to satisfactory levels, more shields can be added. Frequently, cable with the necessary electrical characteristics is available in only one version. If additional shielding is needed, special shielding conduit is available that meets security standards.

Another way to eliminate the cable radiation problem entirely is by using fibre-optic cable. Fibre optics technology uses a glass fibre to carry a beam of light. Information is passed when the light is modulated. With fibre optics no signal is emitted outside the cable; therefore, data cannot be intercepted. Because fibre-optic cable is also extremely difficult to tap into physically, it is ideal for security purposes.

7. CALL-BACK SECURITY

Remote access to networks is a significant threat to data security. Remote workstations are part of many LAN environments, enabling a user to access networks remotely, log into the network, and use the system as if the user were local. Securing this type of access requires special measures.

Call-back security and user management are part of dial-up systems and can be used with remote PC-to-network traffic (see Figure 12.7). With call-back security, when you want to access a computer, you can call into a different number instead of calling in directly. You indicate that you want to access the network, and the security device arranges for a call-back to your location. In other words, the system has embedded within battery-supported memory a complete listing for every allowed user. Included in this file is a ID number that you must punch in when you want to access the file, a telephone number at which you can be reached, and the host systems to which you are allowed access.

This security device also keeps track of user priorities. If all available lines are busy, the device sets up a queue based on the priority of the user. The device will inform the caller regarding queue position. When a line becomes available, the device contacts the user. Therefore, the user never has to get busy signals. The device also keeps accounting information for traffic statistics and call-backs.

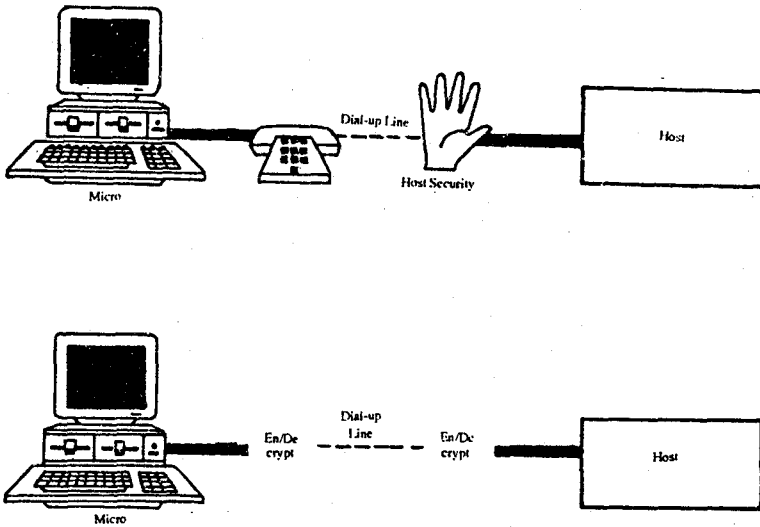


Fig. 12.7 : When a computer system can be accessed through telephone lines, security measures must be taken to protect against unauthorized access. Two ways to protect the system are call-back mechanisms and data encryption.

MANAGEMENT-LEVEL CONCERNS

Managers have a sensitive role in network security (see Figure 12.8). They must help users implement and execute measures like those discussed here and integrate these into network-wide policies that are followed rigorously. These policies also must go beyond the measures that users can implement, but without interfering with users' work.

Managers of sensitive LANs need to address the possibility of their LANs being tapped like telephone lines. With relatively simple electrical devices and a little time, an interloper can tap a LAN cable with little or no immediate evidence. Some LANs can even be tapped from a distance, with devices that monitor the radio-frequency emissions that almost all LANs produce. LANs that permit dial-up connections are particularly susceptible to such taps. LANs based on fibre-optic cable are the most tap resistant.

Call-back modems are a security measure used by many managers of dial-up LAN connections. These modems and their software accept user's calls and then instruct users to enter identifying information and to hang up. The modem then checks the user's access information and calls the user back only after the information is verified. Users who enter information that the system cannot verify are refused network access.

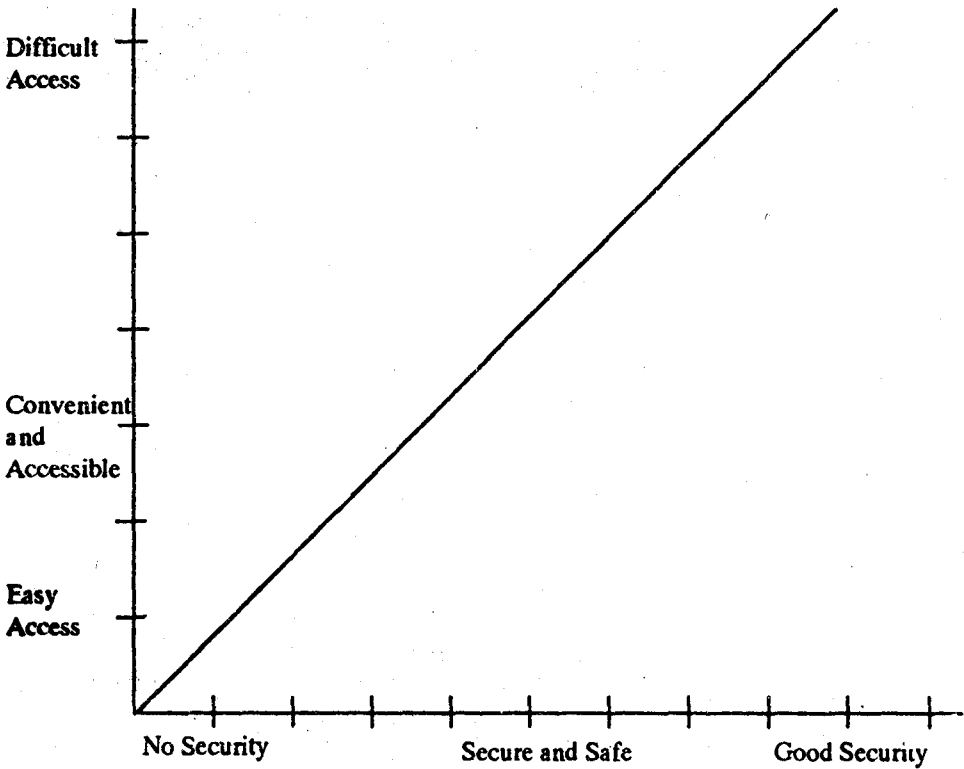


Fig. 12.8 : Good security usually reduces ease of access to the computer. Managers must weigh the trade offs between convenience and security when implementing a specific system

Managers must also monitor connections between their LANs and other networks and computers. Managers must periodically audit access to and from network bridges, routers and other links, and they must regularly update the passwords and other security measures associated with these links. Managers may also have to help users implement more complex personal security measures as their LANs gain access to other networks and systems, and security risks increase.

Managers must also implement measures that provide as much information as possible about network security and about attempted and successful breaches. Ideally, some combination of hardware, software, and physical procedures should be used to provide a near-constant audit of network access and use. This audit will not only help trace the paths of any breaches, but will aid in recovery from any problems these breaches may cause.

Software and procedures that increase accountability can be of great value to a LAN manager and to that manager's organisation. Sufficient information about accountability can limit the liability of an individual, a work group or an organisation, should an accidental or malicious breach of network security result in a loss of tangible assets or in a law-suit for some other reason.

LAN managers are also ultimately responsible for maintaining a constant balance between security measures that are effective, and security measures that interfere with users' work patterns or make users feel that their every move is being monitored. The best way to maintain this balance is to involve users actively and positively in the implementation of any security measures.

Users should also be encouraged to see enhanced security as a way of protecting their own livelihoods and work environments, as well as the assets of their enterprise. The LAN manager has primary responsibility for getting both network users and financial decision makers to see network security as a strategic benefit as well as a basic necessity.

In many cases, the best way to enhance network security is to include security-bolstering procedures and tools alongside aids to other aspects of network operations and management. There are LAN software products which enhance security gracefully. These LAN packages are designed to be invisible to LAN users and can be used to control access to programmes and files, to audit network software for changes, to protect networks against viruses, and to facilitate rapid recovery from disk drive and server failures.

SECURITY IN TCP/IP NETWORKS

Allowing access to your hosts for only the users that you intended is the goal of security in a TCP/IP network. As discussed previously, once a user is logged into a system, the security within that system is all that prevents the user from accessing information that user should not be able to access. Thus, the first level of security in a network is to make sure that all the security on the various hosts themselves is carefully policed. Unfortunately, security in UNIX hosts is based on userid and password. Thus, userids and passwords need to be very difficult to guess. The most effective method used to create a truly secure system is to have userids that are not personal names or initials but computer-generated in some reasonable random style and to have passwords that are likewise computer-generated. Users will not like having userids and passwords that are not easy to remember. But, if someone trespassed into your system, that person will be able to work out userids if it appears that some form of a person's name is being used as a userid. If you find using computer-generated userids too difficult, you should at least force people to change their passwords regularly. Further, you should regularly use password scanning software to make sure people aren't using simple, easy-to-guess passwords like their name or car type.

A second level of security is to ensure that only users who you want are even accessing your network. If your network is completely disconnected from the outside world, you only need to concentrate on the security of your individual hosts. But if you are part of the Internet community and have linked up to the outside world, you need to consider measures to isolate your network from the rest of the Internet. One method is to place a "wall" between your network and the outside networks. This mechanism is often called a firewall because a "fire" in the outside network will not be allowed to enter your network. Routers can be programmed to analyze the network address of a user attempting to access your network and exclude those addresses that are allowed. Thus, routers can be used as firewalls to filter out the network addresses of users who you do not want to access your system.

Other issue is security: The `rlogin` command is sometimes used in place of the `telnet` command because system administrators can set up user validation so that no password is needed for a user to log in on another host. The `telnet` command always requires a password to be entered. Unfortunately, this approach while convenient for users, opens a security hole on the remote system when you use it. With access to the outside world via the Internet a reality for many networks, you should not have any passwordless userids. If you need to provide a *guest* password for a short period of time, you should create a special account for this purpose and then only assign a password when you want to provide access to that guest. By the way, userids on most UNIX systems can be set up not to allow login on that userid at all. These userids are normally present in the system for allowing ownership of system files.

Another issues of security involves permitting the use of anonymous file transfers. You can set up your system so that a file can be transferred between you system so that a file can be transferred between your system and another system without the user having a userid registered on your system. This is accomplished by setting up the FTP user with a special home directory for that user. Then a user can log in using the user anonymous and any password will be accepted for that user. Once logged in, the anonymous user will have access to those files that are in the home directory of the anonymous user. With careful attention to the permissions on other directories, only this one directory can be made available to outside users.

FIREWALL -FRIENDLY FTP

The FTP protocol uses a secondary TCP connection for actual transmission of files. By default, this connection is set up by an active open from the FTP server to the FTP client. However, this scheme does not work well with packet filter-based firewalls, which in general cannot permit incoming calls to random port numbers.

If, on the other hand, clients use the `PASV` command, the data channel will be an outgoing call through the **firewall**. Such calls are more easily handled, and present fewer problems.

An active open is done by the server, from its port 20 to the same port on the client machine as was used for the control connection. The client does a passive open.

For better or worse,, most current FTP clients do not behave that way. A new connection is used for each transfer; to avoid running afoul of TCP's `TIMEWAIT` state, the client picks a new port number each time and sends a `PORT` command announcing that to the server.

If a packet filter is used (as, for example, provided by most modern routers), the data channel requests appear as incoming calls to unknown ports. Most firewalls are constructed to allow incoming calls only to certain believed-to-be-safe ports, such as `SMTP`. The usual

compromise is to block only the "server" area, i.e., port numbers below 1024. But that strategy is risky; dangerous services such as X Windows live at higher-numbered ports. Outgoing calls, on the other hand, present fewer problems, either for the firewall administrator or for the packet filter. Any TCP packet with the ACK bit set cannot be the packet used to initiate a TCP connection; filters can be configured to pass such packets in the outbound direction only. We thus want to change the behavior of FTP so that the data channel is implemented as a call from the client to the server.

Fortunately, the necessary mechanisms already exist in the protocol. If the client sends a PASV command, the server will do a passive TCP open on some random port, and inform the client of the port number. The client can then do an active open to establish the connection.

Recommendation

It is recommended that vendors convert their FTP client programs (including FTP proxy agents such as Gopher) to use PASV instead of PORT. There is no reason not to use it even for non-firewall transfers, and adopting it as standard behavior will make the client more useful in a firewall environment.

SECURITY CONSIDERATIONS

Few people feel that packet filters are dangerous, since they are very hard to configure properly. But they are quite popular. Another common complaint is that permitting arbitrary outgoing calls is dangerous, since it allows free export of sensitive data through a firewall. Some firewalls impose artificial bandwidth limits to discourage this. While a discussion of the merits of this approach is beyond the scope of this memo, we note that the sort of application-level gateway necessary to implement a bandwidth limiter could be implemented just as easily using PASV as with PORT.

Using PASV does enhance the security of gateway machines, since they no longer need to create ports that an outsider might connect to before the real FTP client. More importantly, the protocol between the client host and the firewall can be simplified, if there is no need to specify a "create" operation.

Concerns have been expressed that this use of PASV just trades one problem for another. With it, the FTP server must accept calls to random ports, which could pose an equal problem for its firewall, believe that this is not a serious issue, for several reasons.

First, there are many fewer FTP servers than there are clients. It is possible to secure a small number of special-purpose machines, such as gateways and organizational FTP servers. The firewall's filters can be configured to allow access to just these machines. Further precautions can be taken by modifying the FTP server so that it only uses very high-numbered ports for the data channel. It is comparatively easy to ensure that no

dangerous services live in a given port range. Again, this is feasible because of the small number of servers.

CIRCUIT GATEWAYS

Just as the packet-filtering gateways operate at the network-layer level of a TCP/IP network, a circuit-gateway firewall operates at the transport-layer level--specifically for TCP connections. As discussed, packet-filtering gateways establish an electronic barrier that examines every network packet as the packet passes through the firewall. A circuit gateway, on the other hand, only creates an electronic barrier when two Internet hosts initially establish a TCP connection. When a client program tries to connect to a server on a host that a circuit gateway protects, the circuit-gateway firewall (rather than the server) actually accepts the connection using a special type of relay software. In other words, a circuit-gateway firewall sits as a barrier between both ends of a TCP connection; the gateway's relay software transfers data between the client and server programs on either side of the gateway. As you might suspect, only client programs that know how to talk to the circuit gateway can reach the server on the other side of the firewall. In other words, circuit-gateways require the use of special client programs.

Clients that want access to the server must negotiate a connection with the circuit-gateway relay that let the data transfer occur. The relay software intercepts connection requests that occur at all protocol ports the network security manager defines. After a client successfully negotiates a connection with the relay software, the circuit-gateway becomes essentially invisible. In other words, the circuit-gateway firewall establishes a security barrier only during connection negotiations. After the relay software validates the connection request, data transfer proceeds as though the relay software did not exist--the relay does not examine the content of the packets that pass through the firewall. After the security negotiations complete, the relay software acts much like a wire. In other words, the relay software becomes another part of the transmission medium. Remember, the relay software on the circuit gateway handles the security negotiations which initially occur and then essentially makes itself invisible. In other words, after the connection negotiations complete, the circuit gateway becomes invisible. Obviously, since it's *your* circuit gateway, you will have the special client software necessary to negotiate the connection. The custom client software makes the security negotiations relatively transparent.

CONCLUSIONS

Every expert in home, automobile or business security is quick to point out that there is no lock that cannot be picked, given sufficient time and inclination. The incentive behind sophisticated locks, and policies that encourage and enforce their use, is therefore to make a given facility as difficult and daunting as possible to a potential thief.

Sound network security schemes must accomplish similar goals: They must deter potentially malicious users. In addition, they must encourage users to "lock" their LANs like

they lock their cars and buildings. LAN security strategies must also protect networks from non-malicious accidental incursions, especially by those inexperienced with LANs.

Technology alone is inadequate to ensure security. True network security is a human issue, with responsibility divided between users and managers, just as network processing is increasingly divided between clients and servers. If viewed as a type of client-server risk management, security naturally becomes part of a larger network strategy to ensure total network reliability and to encourage users to participate actively in the protection of their vital network assets.