CHAPTER 6

WIRELESS NETWORK OPERATION

6.1 INTRODUCTION

In the previous chapters, we provided the principles of radio propagation, wireless modem design, wireless access methods, and deployment of cellular systems. These issues were all related to the air interface design and physical characteristics of the wireless medium which are needed to connect a MS to a BS or access point that then connects to the backbone wired networks. To support mobile operation, the backbone network has to add several new functionalities that do not exist for the wired terminal operations, because they are not usually necessary. These functionalities include mobility and location management, radio resource and power management, and security. The very nature of mobile communications implies that the MS is constantly changing locations, warranting a need for tracking the mobile and restructuring existing connections as it moves. Mobility and location management handle the operations required for these purposes. As we have discussed earlier, bandwidth is a scarce resource as also is the battery power of the MS. Also we have seen that a consequence of employing a cellular topology to "multiply" the bandwidth results in wireless networks that are limited by interference. Radio resource and power management schemes are used to address operational aspects related to reducing interference, improving battery life and handling the scarce radio resources. Wireless communications are inherently vulnerable to eavesdropping and need security features that are often not very important in wired networks because of the physical lack of access to the medium. A number of algorithms and methodologies have been implemented in different wireless networks to implement these three features. In this chapter we provide an overview of the techniques that are used for implementation of these three features in wireless networks.

6.2 MOBILITY MANAGEMENT

The primary advantage of wireless communications is the ability to support tetherless access to a variety of services, whether voice oriented as in the case of cellular radio and PCS or data services and access to the Internet as in the case of mobile data networks and wireless LANs. Tetherless access implies that the user has the ability to move around while connected to the network and continuously possesses the ability to access the services provided by the system to which the user is attached. This leads to a variety of issues because of the way in which most communications networks operate. First, in order for any message to reach a particular destination, there must be some knowledge of where the destination is (location) and how to reach the destination (route). In static networks, where the end terminals are fixed, the physical connection (wire or cable) is sufficient to indicate the destination. In wireless networks, where the terminal may be anywhere, there must be a mechanism to locate the terminal in order to deliver the communication to it. *Location management* refers to the activities a wireless network should perform in order to keep track of where the MS is. As discussed in Chapter 5, the most common wireless topology uses multiple cells to provide coverage over a larger area.

The location of the MS must be determined such that there is a knowledge of which point of access (BS or AP) is serving the cell in which the MS is located. Second, once the destination is determined, it is not enough to assume that the destination will remain at the same location with time. When a MS moves away from a BS, the signal level from the current BS degrades, and there is a need to switch communications to another BS. *Handoff* is the mechanism by which an ongoing connection between a MS and a correspondent terminal is transferred from one point of access to the fixed network to another. *Handoff management* handles the messages required to make the changes in the fixed network to handle this change in the location during an ongoing communication. Location and handoff management together are commonly referred to as *mobility management* [AKY98].

6.2.1 Location Management

Location management involves tracking of the location of the MS, as it moves, for delivery of voice or data communication to it. In the case of voice networks, when a call is made to a mobile number, a dedicated channel has to be set up from the calling party to the called party for the conversation to proceed. For this, a circuit has to be set up over the fixed part of the network, and a pair of radio channels have to be allocated to the MS for the voice conversation. The MS has to be located to set up this dedicated channel. Note that this is before the actual conversation takes place. If the MS moves during the course of a conversation, the steps taken to handle the continuity of conversation is called handoff and handoff management. In the case of data networks, packets are addressed to a destination terminal. Routers, within the data network, will use the destination address to deliver the packet. The address information is usually hierarchical and fixed, which means that the address points toward a physical location. If the terminal is fixed, the packet is routed appropriately to the physical location of the terminal. In the case of a MS, some steps are required to determine where it is before the packet is routed to it. Another important functionality of location management is to determine the status of the MS. If the MS is switched off, the network should be aware that it is unreachable so that appropriate action may be taken depending on the service requested. For example, short messages may be stored on a server for later delivery.

Location management in general has three parts to it: location updates, paging, and location information dissemination. *Location updates* are messages sent by the MS regarding its changing points of access to the fixed network. These updates may have varying granularity and frequency. Each time the MS makes an update to its location, a database in the fixed part of the network has to be updated to reflect the new location of the MS. Whether there is a change in the location, the update message will be transmitted over the air and over the part of the fixed network. Because the updates are periodic, there will be some uncertainty in the location of the MS to something around a group of cells. In order to deliver an incoming message to the MS, the network will have to *page* the MS in such a group of cells. The paged terminal will respond through the point of access that is providing coverage in its cell. The response will enable the network to locate the terminal to within the accuracy of the cell in which it is located. Procedures can then be initiated to either deliver the packet or set up a dedicated communications channel for voice

conversation. In order to initiate paging however, the calling party or the incoming message should trigger a location request from some fixed network entity. The fixed network entity will then access some kind of database that will contain the most current location information related to the particular MS and use this information to generate the paging request, as well as deliver the message or set up a channel for the voice call. *Location information dissemination* refers to the procedures that are required to store and distribute the location information related to the MSs serviced by the network.

The basic issue in location management is the trade-off between the cost of the nature, number and frequency of location updates, and the cost of paging [WON00]. If the location updates are too frequent and the incoming messages few, the load on the network becomes an unnecessary cost, both in terms of the usage of the scarce spectrum, as well as network resources for updating and processing of the location updates. If the location updates are few and infrequent, a larger area and thus a larger number of cells will have to be paged in order to locate the MS. Paging in all the cells where the mobile is not located is a waste of resources. Also depending on the way paging is performed, there may be a delay in the response of the MS because the paging in the cell in which it is located might be performed much later than the cell in which it last performed its location update. For applications such as voice calls, this will result in unnecessary call dropping because the MS did not respond in a reasonable time. In the case of data networks, depending on the type of mobility management scheme implemented, packets might simply be dropped if the MS is not located correctly.

As we discussed earlier in this section, location management consists of three activities—location updates, paging, and location information dissemination. There are different types of location management schemes that employ a variety of location update mechanisms, paging schemes, and dissemination architectures. We discuss these in the following sections.

6.2.1.1 Location Update Algorithms

Location update algorithms are usually of two types—static and dynamic [WON00]. In static location updates, the topology of the cellular network decides when the location update needs to be initiated. In dynamic location updates, the mobility of the user, as well as the call patterns, is used in initiating location updates.

In the most common form of static location updates, which is the case in most cellular networks, a group of cells is assigned a *location area* (LA) identifier, as shown in Figure 6.1. Each BS in the LA broadcasts this identification number periodically over some control channel. An MS is required to continually listen to the control channel for the LA identifier. When the identifier changes, the MS will make an update to the location by transmitting a message with the new identifier to the databases containing the location information. If there is an incoming message, paging is performed in the group of cells corresponding to the location identifier stored in the database. The MS usually responds (unless the location area identifier has changed in the meanwhile), and the communication can be delivered successfully.

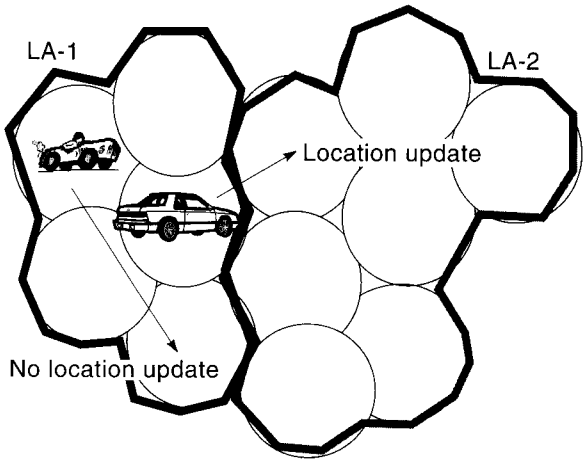


Figure 6.1 Location area (LA) based location updates.

Example 6.1: Location Update Mechanism in GSM

In GSM, an LA identity, also called a paging area, is used for location updates. An LA usually consists of a group of cells controlled by a base station controller (BSC). An MS will perform a location update under three circumstances: (1) Upon powering up, it compares the location area identity it previously had recorded with the one currently being broadcast. If the two location area identities are different, a location update is performed. (2) When the MS crosses the boundary of a location area, it performs a location update. (3) After a period of time predetermined by the network, a location update is performed to ensure that the MS is available. In case (2), the MS detects a change in LA because the BS broadcasts the location area identity, which the MS is required to monitor and compare with the stored value. In case (3), the update mechanism might be costly if the MS does not leave an LA for long periods of time.

The primary problem with the static LA identifier approach is that if an MS is frequently crossing the boundary of two LAs as shown in Figure 6.2, there will be a *Ping-Pong* effect of continually switching between two LAs. A solution to this problem is to employ a dwell timer that persists without a location update for sometime to ensure that the location update is worthwhile. Similar problems and associated algorithms to resolve the problem are encountered during handoff which we will see later on in this chapter.

A variety of other static location update schemes is possible. This includes distance-based—where the location update is performed after crossing a certain number of cells; timer-based—where a location update is performed after a certain time elapses; and variations on these two schemes which take into account signaling load on control channels, and the location and velocity of the MS [WON00].

Examples of dynamic location update schemes are the *state-based* and *user-profile-based* location updates. In the state-based location update scheme, the MS makes a decision on when to perform an update based on its current state information. The state information can include several metrics that include the time

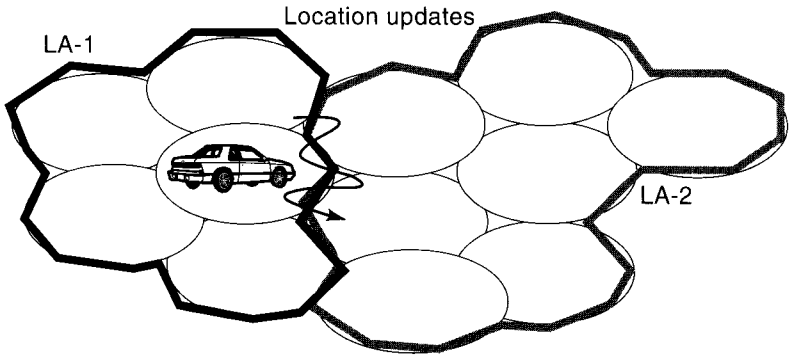


Figure 6.2 Location area ping-pong effect.

elapsed, the distance traveled, the number of LAs crossed, and the number of calls received that could be changed based on the user's mobility and call patterns. The user-profile-based location update schemes maintain a sequential list of LAs that the MS may be located in at different points of time based on the history of the MS. A detailed, comparative evaluation of several of these schemes has been discussed in [WON00].

6.2.1.2 Paging Schemes

Paging is broadcasting a message in a cell or a group of cells to elicit a response from the MS for which a call or message is incoming. Transmitting the page in only the cell in which the mobile terminal is located, which makes the most accurate location estimate, can reduce the cost of paging. The problem with paging in the most accurate cell location is that it is quite impossible to determine location accurately, especially if the location update cost has to be kept low.

Example 6.2: Blanket Paging in GSM

Blanket paging refers to paging the MS in all the cells within an LA simultaneously. This means that, if the LA update is correct, in the very first paging cycle, the MS will receive a paging request and respond to it. The advantage of this system, employed in GSM, is that the delay of the response to paging is kept at a minimum. The disadvantage is that paging has to be done in several cells—all of which have the same LA identity.

Another strategy for paging is to use "closest-cells first" approach. Here the cell where the MS was last seen is paged first followed by subsequent *rings* of cells that are equidistant from this cell in each paging cycle. If there are delay constraints, as in the case of voice calls, several rings may be polled simultaneously in a paging cycle. In general, if the first location estimate is not correct, the next page should be performed so that the probability of locating the mobile is the next largest and so on. The paging is performed in the area corresponding to the last location update, and then subsequent pages are performed in most likely locations

based on parameters that may include past history and distance. A timer is used to declare the MS as unreachable in a particular paging cycle. This is sometimes called *sequential paging*. Results indicate that blanket polling provides the lowest delay at small load, whereas sequential paging can sustain a higher paging request rate, especially when there are several incoming calls to a certain area. The behavior pattern of mobile terminals, as well as the user profile, may also be employed in paging algorithms in a manner similar to location update mechanisms.

6.2.1.3 Location Information Dissemination

When there is an incoming packet to the MS, there is the need for at least one fixed network entity, whose location and address is known, which can be reached to obtain information about the MS. In general, this is often referred to as an *anchor*. The anchor has some information regarding the location and routing information of the MS. If a single anchor is used for all MSs, not only is the load on this entity increased making it a bottleneck for communications, but also it makes a failure point that can result in the collapse of the network. Usually multiple anchor points are employed. What we describe further is in general terms as to how network entities and databases are employed for location, and it can be also extended to other applications such as handoff management. Specific implementations are different.

Every MS is associated with a *home network* and a *home database*. The home database keeps track of the profile of the MS—such as the mobile identification, authentication keys, subscriber profile, accounting, and location. The location of the mobile is maintained in terms of a *visiting network,* where the MS is located, and a *visiting database,* which keeps track of the MSs in its service area. The home and visiting databases communicate with each other to authenticate and update each other about the MS. We will see more of this in the section on handoff management.

Example 6.3: Location Information Dissemination in GSM

In GSM, the home and visiting databases are called *home location register* (HLR) and *visiting location register* (VLR), respectively. When the MS observes a change in the LA identity, it transmits a location update message through the BS to a MSC. The MSC contacts its VLR with the location update. The VLR does nothing if it serves both the old and new LA. If the VLR has no information about the mobile terminal, it contacts the HLR of the MS via a location registration message. The HLR authenticates and acknowledges the location registration, updates its own database, and sends a message to the old VLR to cancel the registration there.

Example 6.4: Call Delivery in GSM

When a call is made to a mobile telephone number, the *anchor* entity contacts the MSC associated with the HLR of the mobile terminal. The HLR contacts the VLR that is associated with the MS and enables call setup. A detailed description of the call delivery in GSM is presented in Chapter 7.

6.2.1.4 Emerging Issues in Location Management

Location management has several potential issues associated with it [AKY98], [WON00]. Primarily, there is a lot of research going on in database architectures for next-generation wireless networks. Access to the database and management of queries is very important in order to reduce delay and maintain quality. To reduce the load on a centralized database (such as an HLR), local caches of the mobile terminal information can be maintained. Similar strategies are being considered in Mobile IP, as we will see later on. Alternative location update strategies and paging algorithms are being investigated. An important factor that influences the performance of all these techniques is traffic modeling, which can accurately represent the nature of incoming calls, paging requests, and movement of the MS.

6.2.2 Handoff Management

Handoff management involves the entire gamut of issues and actions that are required to handle an ongoing connection when a mobile terminal moves from the coverage of one point of access to another. Handoff [POL96], [TRI97], [TRI98] is extremely important in any mobile network because of the default cellular architecture employed to maximize spectrum utilization. Handoff, in the case of cellular telephony, involves the transfer of a voice call from one BS to another. In the case of WLANs, it involves transferring the connection from one AP to another. In hybrid networks, it will involve the transfer of a connection from a BS to another, from an AP to another, between a BS and an AP, or vice versa.

For a voice user, handoff results in an audible click interrupting the conversation for each handoff [POL96], and because of handoff, data users may lose packets and unnecessary congestion control measures may come into play [CAC95]. Degradation of the signal level is, however, a random process, and simple decision mechanism such as those based on signal strength measurements result in the *Ping-Pong effect*. The Ping-Pong effect refers to several handoffs that occur back and forth between two BSs. This has a severe toll on both the user's quality perception and the network load. A way of eliminating the Ping-Pong effect is to persist with a BS for as long as possible. However, if handoff is delayed, weak signal reception persists unnecessarily, resulting in a lower voice quality, increasing the probability of call drops and/or degradation of QoS. Consequently, more complex algorithms are needed to decide on the optimal time for handoff. Handoff also involves a sequence of events in the backbone network that include rerouting the connection and reregistering to the new point of access, which are additional loads on the network traffic. Handoff has an impact on traffic matching and traffic density for individual BSs (because the load on the air-interface is transferred from one point of access to another). In the case of random access techniques employed to access the air interface, or in the case of CDMA, moving from one cell to another impacts QoS in both cells because throughput and interference depend on the number of terminals competing for the available bandwidth.

Although significant work has been done on handoff mechanism in circuit switched mobile networks [POL96], [TRI98], there is not much of literature available for packet switched mobile networks. Performance measures such as call blocking and call dropping probabilities are applicable only to real-time traffic and

may not be suitable for bursty traffic which exists in client-server type of applications. When a voice call is in progress, allowed latency is very limited and resource allocation has to be guaranteed, and although occasionally some packets may be dropped and moderate error rates are permissible, retransmissions are not possible, and connectivity has to be maintained continuously. On the other hand, bursty data traffic by definition needs only intermittent connectivity and can tolerate greater latencies and employ retransmission of lost packets. In such networks handoff is warranted only when the terminal moves out of coverage of the current point of attachment or the traffic load is so high that a handoff may result in greater throughput and utilization.

There are a variety of issues related to handoff. In particular we can consider handoff as consisting of two different steps as shown in Figure 6.3. In the first step, the handoff management process determines that a handoff is required (handoff decision and initiation). In the second step, the rest of the network is made aware of the handoff, and the connection is restructured to reflect the new location of the MS. Note that there is an *anchor* in the fixed part of the network that must be involved in the handoff management process in a manner similar to location management. Several issues arise during the handoff management process.

As shown in Figure 6.4, these issues are divided into two categories: architectural issues and handoff decision time algorithms. Architectural issues are those related to the methodology, control, and software/hardware elements involved in rerouting the connection. Issues related to the handoff decision time algorithms are the types of algorithms, metrics used by the algorithms, and performance evaluation methodologies.

6.2.2.1 Architectural Issues in Handoff

Handoff procedures involve a set of protocols to notify all the related entities of a particular connection that a handoff has been executed and that the connection has to be redefined. In data networks, the MS is usually registered with a particular point of attachment. In voice networks, an idle MS would have selected a

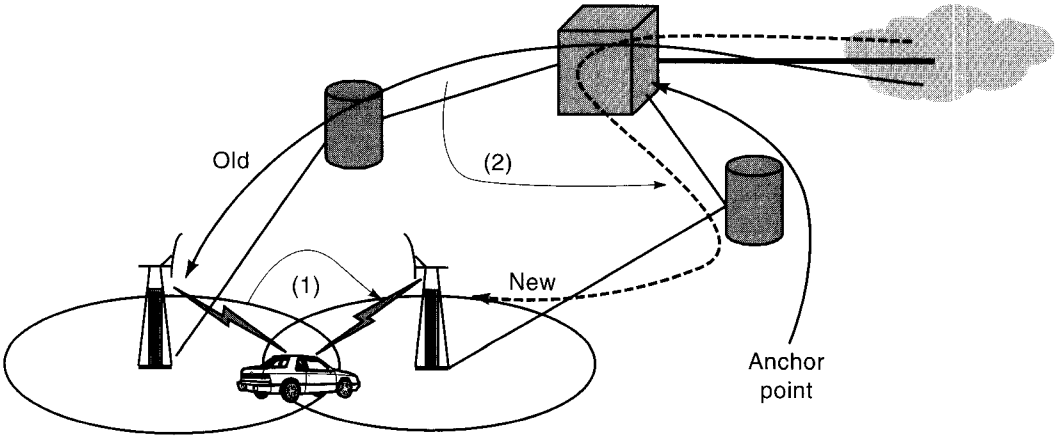


Figure 6.3 Two basic actions during handoff.

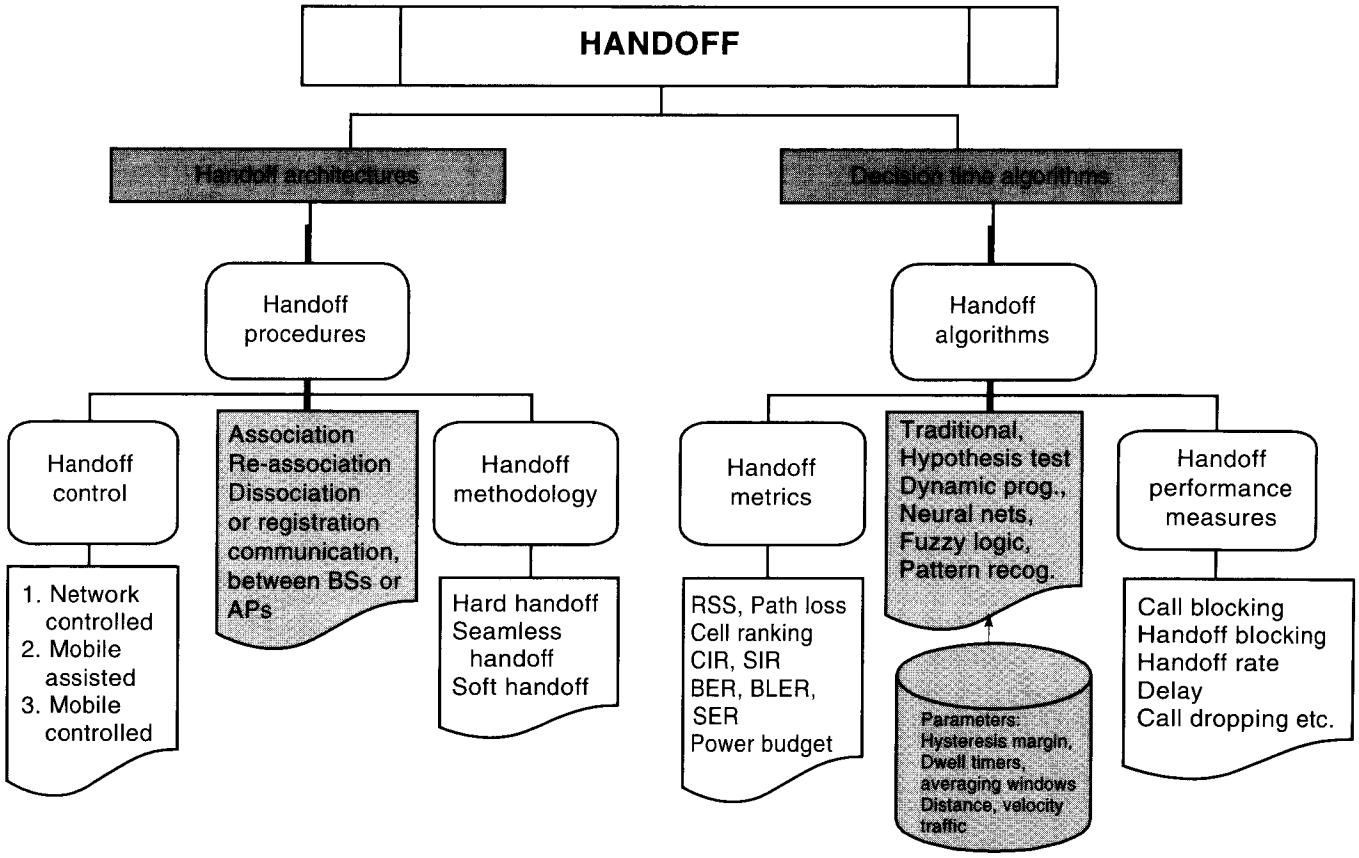


Figure 6.4 Important issues involved in the handoff mechanism.

particular BS that is serving the cell in which it is located. This is for the purpose of routing incoming data packets or voice calls appropriately. When the MS moves and executes a handoff from one point of attachment to another, the old serving point of attachment has to be informed about the change. This is usually called dissociation. The MS will also have to reassociate itself with the new point of access to the fixed network. Other network entities that are involved in routing data packets to the MS or in switching voice calls have to be aware of the handoff in order to seamlessly continue the ongoing connection or call.

Example 6.5: Registration in CDPD

The mobile terminal in CDPD is called the *mobile end system* (M-ES). It announces its presence to the network by sending a message called End_System_Hello to a mobile data intermediate system—MD-IS (which is the network entity controlling a group of cells). Registration is completed when the MD-IS responds with a confirmation. The MD-IS serving the cell has to contact the *home* MD-IS (which controls the home database) to authenticate the M-ES and also update it about the location of the M-ES.

Depending on whether a new connection is created before breaking the old one, handoffs are classified into a hard and seamless handoff. Hard handoff occurs when the mobile terminal completely breaks connection with the old BS before connecting to the new BS and synchronizing itself to it. Seamless handoff refers to the case where the mobile terminal sets up a traffic channel with the new BS before breaking off from the old BS. However, communication is possible only through one BS at a time. In CDMA, the existence of two simultaneous connections during handoff results in soft handoff [TEK91]. Soft handoff is discussed in more detail in Chapter 8.

The decision mechanism or *handoff control* could be located in a network entity (as in cellular voice) or in the mobile terminal (as in mobile data and WLANs) itself. These cases are called network controlled handoff (NCHO) and mobile controlled handoff (MCHO), respectively. In GPRS, information sent by the mobile terminal can be employed by the network entity in making the handoff decision. This is called the MAHO.

Example 6.6: Handoff Control in Different Systems

In AMPS, the analog 1G standard, handoff decision is NCHO. The mobile telephone switching office uses the RSS measurements from an MS at different BSs to initiate handoff. In the case of IEEE 802.11 LANs, the mobile station controls handoff decision (MCHO). It monitors the *beacon* of several APs to decide which AP to connect to. The network has no role in deciding when to make a handoff.

In any case, the entity that decides on the handoff uses some metrics, algorithms, and performance measures in making the decision. The measurement and handoff decision are usually part of the radio resource management procedures. However, we look at handoff decision mechanisms in this section to keep the procedures for handoff together.

6.2.2.2 Handoff Decision Time Algorithms

Several algorithms are being employed or investigated to make the correct decision to handoff [POL96], [TRI98]. Traditional algorithms employ thresholds to compare the values of *metrics* from different points of attachment and then decide on when to make the handoff. A variety of metrics have been employed in mobile voice and data networks to decide on a handoff.

Primarily, the RSS measurements from the serving point of attachment and neighbouring points of attachment are used in most of these networks. Alternatively or in conjunction, the path loss, carrier-to-interference ratio (CIR), signal-to-interference ratio (SIR), BER, block error rate (BLER), symbol error rate (SER), power budgets, and cell ranking have been employed as metrics in certain mobile voice and data networks. In order to avoid the Ping-Pong effect, additional parameters are employed by the algorithms such as hysteresis margin, dwell timers, and averaging windows. Additional parameters (when available) may be employed to make more intelligent decisions. Some of these parameters also include the distance between the MH and the point of attachment, the velocity of the MH, traffic characteristics in the serving cell, and so on.

Traditional handoff algorithms are all based on the RSS or received power P. Some of the traditional algorithms [POL96] are as follows:

1. *RSS:* The BS whose signal is being received with the largest strength is selected (choose BS B_{new} if $P_{new} > P_{old}$).
2. *RSS plus Threshold:* A handoff is made if the RSS of a new BS exceeds that of the current one and the signal strength of the current BS is below a threshold T (choose B_{new} if $P_{new} > P_{old}$ and $P_{old} < T$).
3. *RSS plus Hysteresis:* A handoff is made if the RSS of a new BS is greater than that of the old BS by a hysteresis margin H (choose B_{new} if $P_{new} > P_{old} + H$).
4. *RSS, Hysteresis, and Threshold:* A handoff is made if the received signal strength of a new BS exceeds that of the current one by a hysteresis margin H and the signal strength of the current BS is below a threshold T (choose B_{new} if $P_{new} > P_{old} + H$ and $P_{old} < T$).
5. *Algorithm plus Dwell Timer:* Sometimes a dwell timer is used with the other algorithms. A timer is started at the instant when the condition in the algorithm is true. If the condition continues to be true until the timer expires, a handoff is performed.

Figure 6.5 illustrates these algorithms in the case of a mobile terminal traveling between two BSs along a straight line. Note that the RSS is not smooth as shown in this figure but more random as illustrated in Figure 6.6.

Recently other techniques are emerging such as hypothesis testing [LIO94], dynamic programming [REZ95], and pattern recognition techniques [COX96] based on neural networks or fuzzy logic systems [TRI97] (for an excellent survey of various algorithms, see [POL96], [TRI97]). These complicated algorithms are necessitated by the complexity of the handoff problem, especially in hybrid data or voice networks. The mobile terminal has to monitor the air for wireless data services that may be available for attachment. As an example, consider an MS that

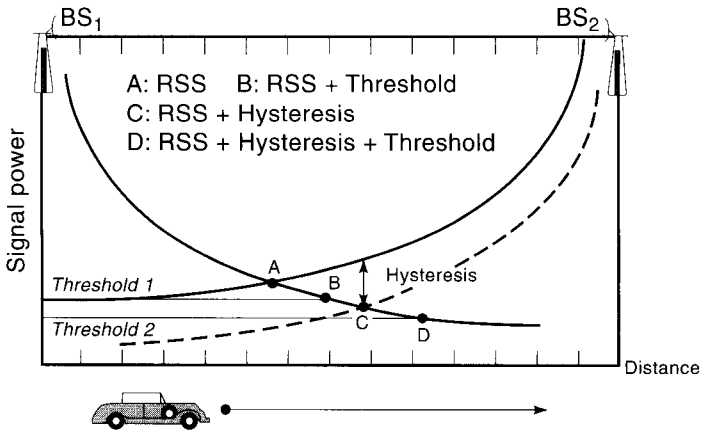


Figure 6.5 Traditional handoff algorithms using RSS thresholds and hysteresis.

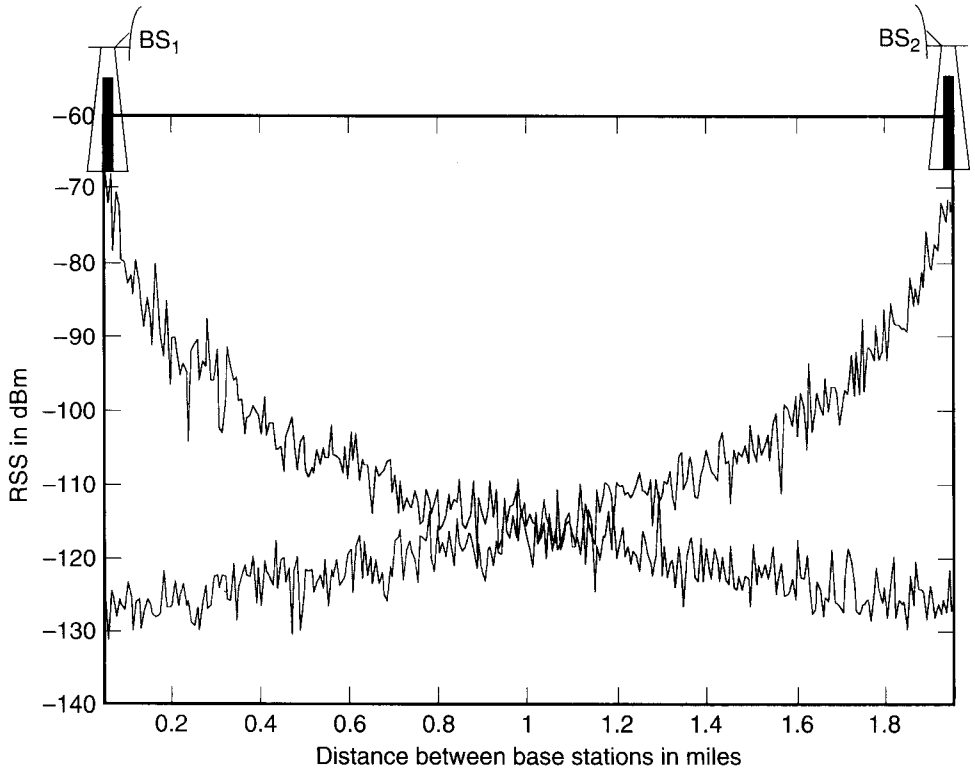


Figure 6.6 Sample RSS from two base stations as seen by an MS traveling in a straight line between them.

could connect to either an 802.11 WLAN AP connected to a LAN or a GPRS BS subsystem (BSS) connected to a backbone GPRS network. There must be a mechanism or algorithm within the mobile terminal that will enable it to choose the best available service and switch to this service as soon as it is available. For example, the mobile terminal must be able to switch from the GPRS service to the WLAN AP as soon as it detects the availability of a connection to an access point. Most of the emergent algorithms are in their nascent stages, and they have been analyzed or simulated only for voice networks and only for extremely simple scenarios.

The performance of handoff algorithms is determined by their effect on certain performance measures. Most of the performance measures that have been considered such as call blocking probability, handoff blocking probability, delay between handoff request and execution, call dropping probability, and so on are related to voice connections. Handoff rate (number of handoffs per unit time) is related to the ping-pong effect, and algorithms are usually designed to minimize the number of unnecessary handoffs. Although minimizing the handoff rate is important in mobile data networks, other issues include throughput maximization and maintaining QoS guarantees during and after handoff. However, these issues have not received sufficient attention in the literature.

6.2.2.3 Generic Handoff Management Process

In this section, we show the different messages and processes that are required for handoff management in a generic wireless network. As in the case of location management, the specific implementations will be different. We consider some specifics in later chapters.

In Figure 6.7, a generic architecture is shown for the handoff management process. There are two types of databases in the network; the home database that also acts as the anchor and the visiting database. Every mobile terminal is registered with a home database that keeps track of the profile of the mobile terminal.

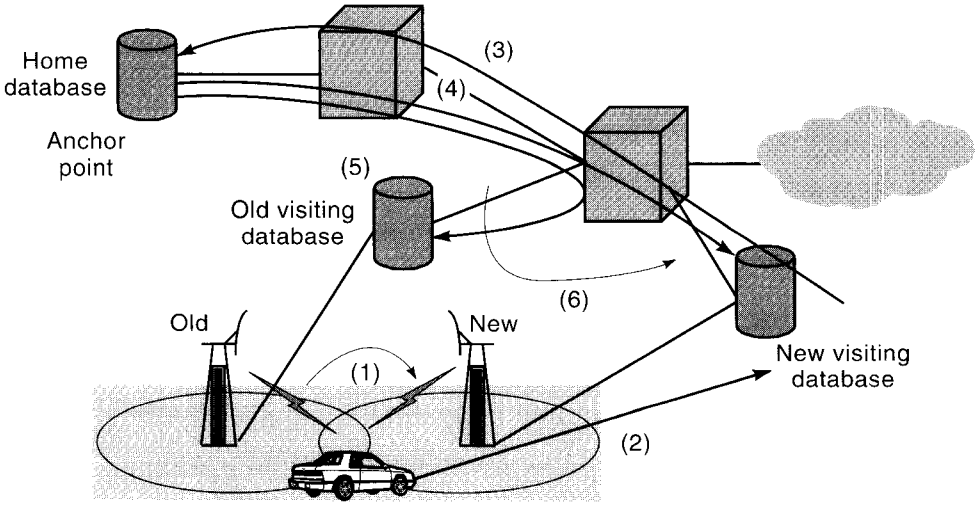


Figure 6.7 Generic handoff management process.

The visiting database keeps track of the mobile terminals in its service area. The home and visiting databases communicate with each other during the handoff management process.

1. In the first step, a decision is made to handoff, and handoff is initiated. This decision, as discussed earlier, may be made in the network by some entity with or without the help of the mobile terminal, or at the mobile terminal. For this purpose the decision time algorithms are employed.
2. The mobile terminal registers with the "new" visiting database via a handoff announcement message. This is the first information to a network entity in the case of a mobile-controlled handoff. In the case of a network controlled or mobile assisted handoff, the new visiting database may already be aware or expecting this message.
3. The new visiting database communicates with the home database to obtain subscriber profile and for authentication. This is the first information exchange between network entities about the changed location of the mobile in MCHO. In the case of MAHO or NCHO, these entities may already be in communication.
4. The home database responds to the new visiting database with the authentication of the mobile. If the mobile is authenticated, in the case of circuit-switched connections, a pair of traffic channels that might be kept ready is allocated to the mobile terminal for continuing the conversation. In the case of packet data traffic, no such dedicated channels are required because the traffic is bursty. The two databases are updated for delivering new messages that may arrive to the mobile terminal. The new visiting database includes the mobile terminal in its list of terminals that are being serviced by it.
5. The home database sends a message to the old visiting database to flush packets intended for and registration information related to the mobile terminal. This is because packets that may have been routed to the old visited network while the mobile was making a handoff need to be dropped or redirected, and the old visiting database needs to clear resources it had maintained for the mobile terminal because they are no longer required.
6. The old visiting database flushes or redirects packets to the new visiting database and removes the mobile terminal from its list.

Each of these steps is important in order to correctly, securely, and efficiently implement handoff and release resources that are otherwise not used in the system.

Mobility management procedures have details that are specific to the respective systems. They need some description of the network entities because the nomenclature for the databases and the controlling entities are different with different functionality. Descriptions of mobility management procedures are provided in subsequent chapters where individual technologies are described. There are other architectural issues such as handoff between channels in a cell (intracell handoff), handoff between two BSs associated with the same database (intercell intradomain handoff), and different databases (intercell interdomain handoff), and so on. These issues are also discussed when we describe specific technologies later

on. Also procedures are adopted when an MS returns to a cell from which it had been handed off to simplify the connections in the backbone.

In the following section, we consider mobile IP as a specific handoff management scheme that has two drawbacks—it does not specify either the first step (handoff decision and initiation) or the last step (flushing and redirecting data). These are technology specific as far as mobile IP is considered.

6.2.3 Mobile IP

IP, which is the most popular network layer protocol for data networks, was not designed with wireless or mobile networks in view. Mobile IP tries to address this issue by creating an "anchor" for a mobile host that takes care of packet forwarding and location management. Mobile IP [PER97] is simplified because IP packets do not need mechanisms to set up dedicated bandwidth or channels as in the case of circuit switched connections. However, it solves a different kind of problem that IP created when the terminals were mobile. The IP address is used for dual purposes—for routing packets through the Internet and also as an endpoint identifier for applications in end-hosts. The connections in an IP network use *sockets* to communicate between clients and servers. A socket consists of the following tuple: `<source IP address, source port, destination IP address, destination port>`. A TCP connection cannot survive any address change because it relies on the socket to determine a connection. However, when a terminal moves from one network to another, its address changes. This is because the Internet uses domain names that are converted to an IP addresses. A packet addressed to one IP address gets routed to the *same place* always because the IP address also points to the location of a physical network.

An IP-mobility working group of the Internet Engineering Task Force (IETF) is in charge of activities related to Mobile IP. Several standards and requests for comments (RFCs) related to mobile IP are available [IETweb]. The basic design criteria for mobile IP were (a) compatibility with existing network protocols, (b) transparency to higher layers (TCP through application) and to the user, (c) scalability and efficiency in terms of not requiring a great deal of additional traffic or network elements, and (d) security due to changing locations of the mobile node (MN).

An MN is a terminal that can change its location and thus its point of attachment. The partner for communication is called the *correspondent node* (CN) that can be either a fixed or a mobile node. The IP network where the MN resides is called the *home network*, and the IP network where the MN is visiting is called the *foreign network*. The *home address* of an MN is a long-term IP address assigned to the MN that is part of the home IP network. It remains unchanged regardless of where the MN is, and it is used for domain name system (DNS) determination of the MN's IP address. The *care-of address* (COA) is an IP address in the foreign network that is the reference pointer to the MN when it is visiting the foreign network. The *home agent* (HA) is the anchor in the home network for the MN. All packets addressed to the MN reach the HA first unless the MN is already in its home network. A *foreign agent* (FA) (only in the case of IPv4) acts as the reference point in the foreign network for the MN. The COA is usually the IP address of the FA. The MN can act as its own FA, in which case it is called a *colocated* COA.

6.2.3.1 Location Management in Mobile IP

Location management in Mobile IP is achieved via a registration process and the so-called agent advertisement. Foreign agents and home agents periodically "advertise" their presence using *agent advertisement* messages. The same agent may act as both an HA and an FA mobility extension to ICMP messages which are used for agent advertisements. The messages contain information about the COA associated with the FA, whether the agent is busy, whether minimal encapsulation is permitted, whether registration is mandatory, and so on. The agent advertisement packet is a broadcast message on the link. If the MN gets an advertisement from its HA, it *must* deregister its COAs and enable a gratuitous address resolution protocol (ARP). If an MN does not "hear" any advertisement, it must solicit an agent advertisement using ICMP. The entire connection search flow is shown in Figure 6.8.

Once an agent is discovered, the MN performs either a registration or deregistration with the HA, depending on whether the discovered agent is an HA or an FA. The MN sends a *registration request,* using UDP to the HA through the FA (or directly if it is a colocated COA). The HA creates a *mobility binding* between the MN's home address and the current COA that has a fixed lifetime. The MN should reregister before the expiration of the binding. A *registration reply* indicates whether the registration is successful. A rejection is possible by either the HA or FA for such reasons as insufficient resources, the HA is unreachable, there are too many simultaneous bindings, for failed authentication, and so on. If an MN does

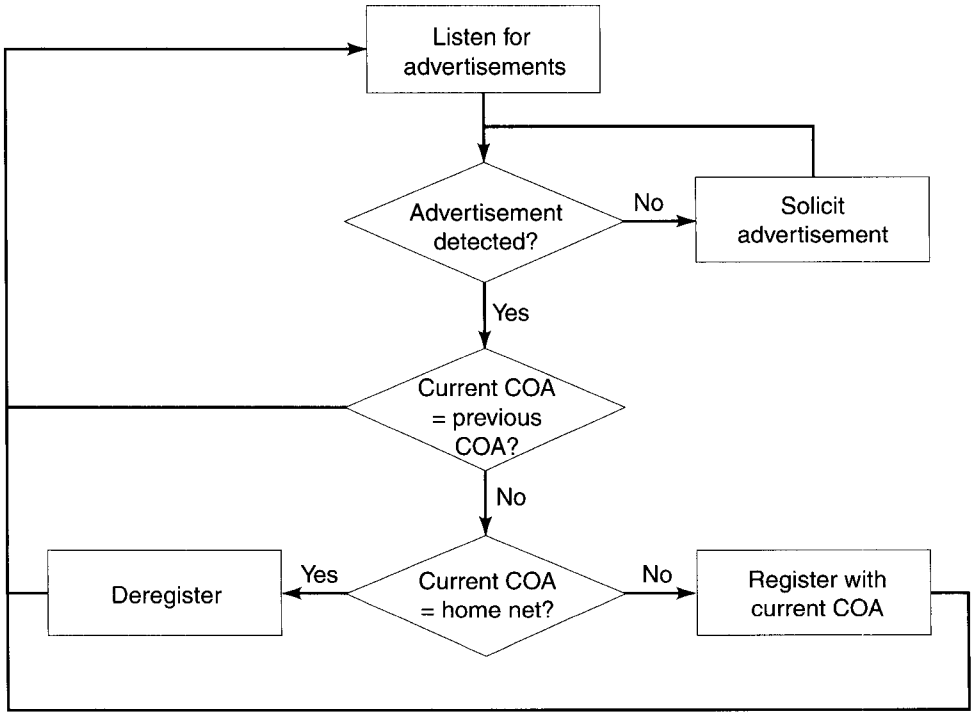


Figure 6.8 Agent discovery procedure.

not know the HA address, it will send a broadcast registration request to its home network called a *directed broadcast*. The response to this request is a reject by every valid HA. The MN uses one of the HA addresses in the reject message to make a valid registration request. The HA and FA maintain lists of MNs in what we can relate to as the home and visiting databases. Upon a valid registration, the HA creates an entry for an MN that has the MN's care of address, an identification field, and the remaining lifetime of the registration. Each foreign agent maintains a visitor list containing the following information: link layer address of the MN, MN's home IP address, UDP registration request source port, HA IP address, an identification field, the registration lifetime, and the remaining lifetime of pending or current registration.

6.2.3.2 Handoff Management in Mobile IP

Mobile IP enables datagrams addressed to the MN at the home address to be delivered wherever the MN is. As shown in Figure 6.9, the CN transmits a datagram to the MN that is routed to MH's home network as usual in step (1). The HA intercepts the packet, encapsulates and tunnels it to FA in step (2). The FA decapsulates and forwards the packet to the MN in step (3). Packets from the MN to the CN are sent as usual (4). This procedure is called triangle routing.

In order to intercept packets addressed to the MN, the HA performs a proxy ARP on behalf of the MN when it is away. The way ARP works is as follows. An ARP request is a broadcast message seeking the MAC (physical) address of a terminal given its IP address. When a packet arrives to the MN, an ARP request is made to obtain its MAC address on the home network. If the MN is away, the HA will respond with its own MAC address. When it returns to the home network, the MN will perform a gratuitous ARP that is an unsolicited ARP reply broadcast to each node on the home network clearing the ARP caches. Forwarding packets is achieved by encapsulation (tunneling). A virtual pipe between tunnel entry point (HA) and tunnel termination point (FA) is created through a datagram that includes the packet from the CN as its payload. The mandatory implementation for Mobile IP is IP-in-IP encapsulation as shown in Figure 6.10 though more efficient implementations (called minimal encapsulation) are optional. As far as the IP

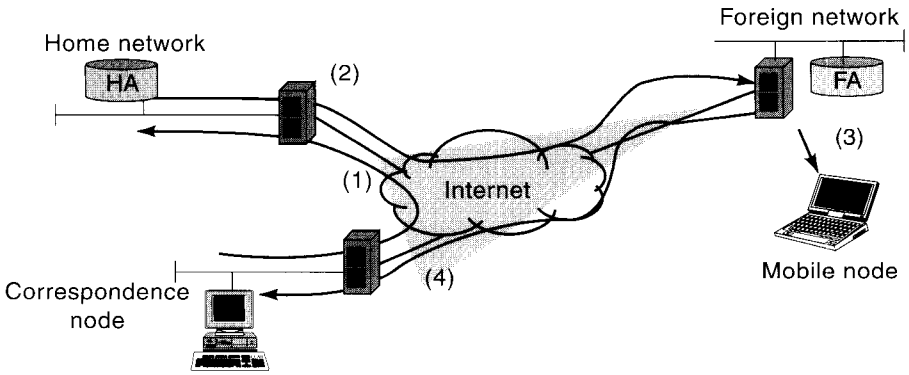


Figure 6.9 Triangle routing in Mobile IP.

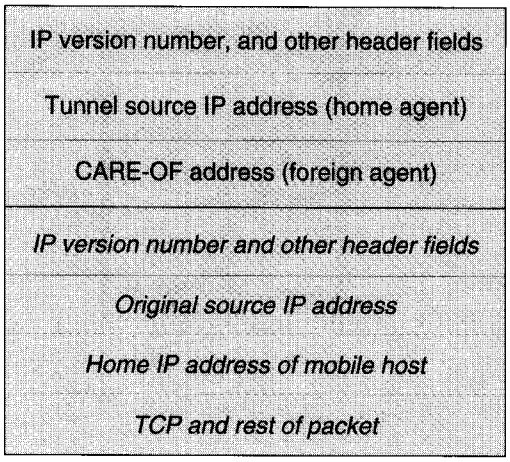


Figure 6.10 IP-in-IP encapsulation.

packet from the correspondent node is concerned, it looks like a single hop within the Internet.

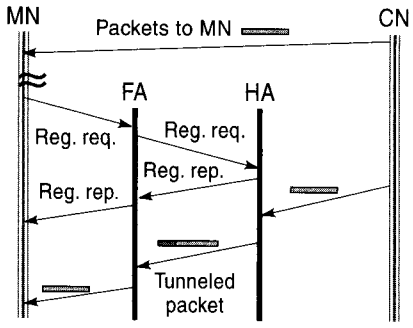
Figure 6.11(a) shows the sequence of events when an MN moves from the home network to a foreign network, and Figure 6.11(b) shows the sequence of events when it returns to the home network.

6.2.3.3 Other Issues in Mobile IP

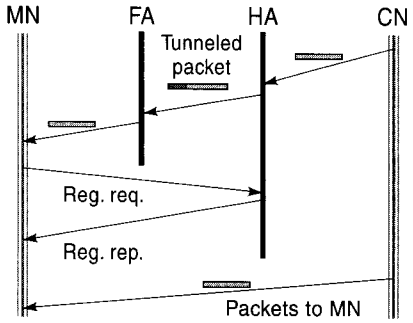
There are several issues in Mobile IP that are under consideration. Because the HA has to tunnel packets, it could be a potential bottleneck in the case of heavy traffic. Triangle routing is inefficient, especially if packets are routed to the home network, only to be tunneled back to a point close to the CN. A solution for both these problems is to enable routers in the Internet to cache the mobility binding and route packets accordingly. The packets addressed to the MN can be detected when they are being routed as packets that need tunneling to a new address and can be routed as such. This, however, leads to issues related to security and the need to change the way in which routers operate.

Suppose an MN changes its foreign network and while a new registration request is in progress, data are being tunneled to the old FA. These data have to be resent by the CN, as the old FA will drop the packets addressed to the MN. After the CN sends the data, the retransmitted data have to be tunneled again. If the old FA can tunnel packets it receives to the new FA, this can reduce delay and congestion. Such a procedure is called *smooth handoff*. It is also possible that the old FA retunnels the packet back to the HA, in what is called a "special tunnel." This enables the HA to detect a "loop" if a new registration request has not been enabled.

Sometimes packets will have to be tunneled through the HA. Two common reasons for this are that firewalls drop outgoing packets that have an IP address which corresponds to another network, so the packet cannot be directly sent from the MN to the CN. Also packets addressed by the MN to hosts on the home network usually have a small time-to-live because they are supposed to be on the same network. A small time-to-live implies that the packet needs to sense the Internet as



(a)



Reg. req.: Registration request
Reg. rep.: Registration response

(b)

Figure 6.11 Sequence of events: (a) when the MN moves to a foreign network and (b) when the MN returns to the home network.

one hop. In both cases, the MN can tunnel the packet to the home agent for retransmission; this procedure is called *reverse tunneling.*

6.3 RADIO RESOURCES AND POWER MANAGEMENT

Radio resource and power management is an important part of any wireless network because of several reasons as explained in the following. The scarcity of radio spectrum has resulted in frequency reuse as discussed in Chapter 5 to increase capacity in a serving area. The installation of multiple BSs to provide service results in certain phenomena that need to be correctly addressed for proper operation of the wireless network. First, signals from MSs operating in the coverage area of one BS cause interference to the signals of MSs operating in the coverage area of another BS on the

uplink. There is a need to reduce such an interference by properly controlling the transmit powers of MSs. Similarly, signals transmitted by one BS will interfere with the signals transmitted by another BS on the downlink. The transmit powers of BSs on interfering channels need to be controlled to minimize this interference. Second, correctly controlling the transmit powers of MSs can enhance their battery life and make mobile terminals lighter and handier to use. Because wireless terminals are mobile, they run on battery power which needs to be conserved as long as possible to avoid the inconvenience of requiring a fixed power outlet for recharging. Most of this power is consumed during transmission of signals. Consequently, the transmit power of the MS must be made as small as possible. In turn, this requires reducing the coverage area of a point of access—whether a BS or an AP, so that the received signals are of adequate quality. Also, as mobile terminals move, the ability to communicate with the current BS degrades, and they will need to switch their connection to a neighboring BS. At some point during the movement, a decision has to be made to handoff from one BS to another. This decision will have to be made based on the expected future signal characteristics from several BSs that may be potential candidates for handoff, the capacities and available radio resources of such BSs, and interference considerations. For example, if an MS continues to communicate with a BS when it is deep into the coverage area of another BS, it will cause significant interference in some other cell that employs the same channel. Third, then, there is a need for the wireless network to keep track of the radio resources, signal strengths, and other associated information related to communication between an MS and the current and neighboring BSs. All these tasks are not undertaken by a single entity. We have already discussed the last issue in the section on handoff decision algorithms. As we discuss in this chapter, several schemes and technologies are employed for the first two issues of radio resources and power management in wireless networks.

We distinguish between power control, power saving mechanisms, energy efficiency, and radio resource management. By *power control*, we shall mean the algorithms, protocols, and techniques that are employed in a wireless network to dynamically adjust the transmit power of either the MS or BS for reducing cochannel interference, near-far interference in the case of CDMA, or other reasons. *Power-saving mechanisms* are employed to save the battery life of a mobile terminal by explicitly making the MS enter a suspended or semisuspended mode of operation with limited capabilities. This is, however, done in cooperation with the network, so as to *not* disrupt normal communications or provide the user with a perception of such a disruption even if there was one. *Energy efficient design* is a new area of research that is investigating approaches to save the battery life, of an MS in fundamental ways such as in protocol design, via coding and modulation schemes, and in software. *Radio resource management* refers to the control signaling and associated protocols employed to keep track of relationships between signal strength, available radio channels, and so on in a system so as to enable an MS or the network to optimally select the best radio resources for communications.

6.3.1 Power Control

In this section we discuss basic power control mechanisms and why they are important through example implementations in cellular networks.

6.3.1.1 Basic Idea in Cellular Networks

Power control has been an issue of importance since the very first deployment of analog cellular systems. As discussed in Chapter 5, cochannel interference limits the capacity of a cellular network. Cochannel interference also causes the quality of a voice signal to deteriorate, and an attempt has to be always made to ensure that cochannel interference is minimized. This translates into forcing a mobile terminal or a BS to operate at the *lowest possible SIR* so that the voice or communications quality is acceptable. This appears to be a paradox because one would expect that it is important to maintain a high SIR for good communications quality. Although this is true in ordinary communications systems, in wireless communications with a cellular topology, operating at a high SIR implies that the transmit power of an MS or a BS is large. A large transmit power in one frequency channel in one cell results in a large cochannel interference in all the closest cochannel cells that employ the same frequency channel, albeit at a sufficient distance away from the given BS. This will reduce the communications quality all around and is not desirable.

Example 6.7: Minimum S_r Operation in AMPS on the Reverse Link

Consider an AMPS network. As discussed in Chapter 5, usually a cluster of seven cells uses the entire spectrum allocated to an operator. The spectrum is then reused in neighboring clusters. The approximate distance between the centers of cochannel cells D_L is $4.58R$ where R is the radius of a cell. Consider an MS in one of the cells of the cluster. Suppose it is located at a distance of $R/2$ from its own BS. Without power control, it would transmit at some power P_t, which is independent of distance. If radio propagation in the cell has a distance-power gradient of four, in order for it to operate such that the BS receives the signal at the lowest possible S_r, the transmit power must be reduced to $P_t/16$ because the BS would otherwise receive a signal that is 16 times stronger. This will in turn reduce the interference it causes to cochannel cells as well as adjacent channel cells. By reducing the transmit power, the mobile terminal will also save on its battery life.

Example 6.8: Effect of Large Transmit Power on the Forward Link in AMPS

Consider an AMPS network with a reuse factor of $N = 7$. Assume that channels 1, 8, 15, and so on are allocated to cells labeled A in Figure 6.12 (see Example 5.8). Channel 1 corresponds to the frequency band 869.0–869.030 MHz. Let the BS in the shaded cell transmit signal on Channel 1 at a transmit power six times as large as the other BSs of its co-channel cells. This is an increase in the transmit power by less than 8 dB. The effect on the SIR observed by the mobile terminal in Figure 6.12 is as follows:

$$S_r \approx \frac{P_t R^{-4}}{5P_t D_L^{-4} + 6P_t D_L^{-4}} = \frac{1}{11}\left(\frac{D_L}{R}\right)^4 \tag{6.1}$$

From Chapter 5, we know that the ratio D_L/R is 4.58 for the case where $N = 7$. Here D_L is approximately the distance of the mobile terminal from its cochannel cells. The SIR given by (1) is around 16 dB that is 2 dB lower than the required value for good communications quality. If all the BSs in the area are erratic, the signals received by the mobile terminals will all be of poor quality.

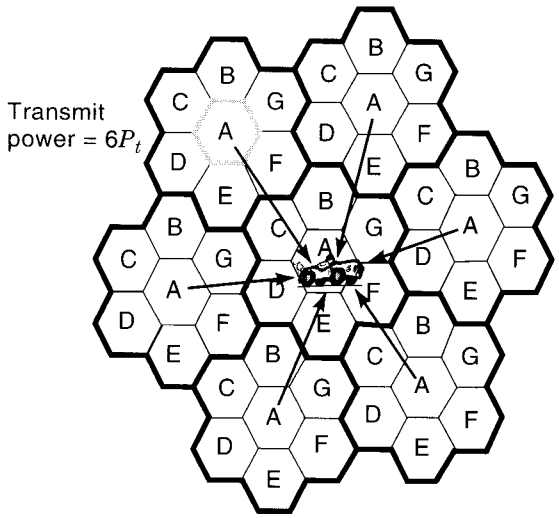


Figure 6.12 Effect of large transmit power.

From these examples, we can see that controlling the transmit power of both MSs and the BSs is important. When power control is applied properly, it can improve the quality of communications by increasing the SIR. An alternative way to view this is in terms of increase in system capacity. If the SIR can be increased, this will imply that a lower frequency reuse can be employed. As discussed in Chapter 5, this will increase the capacity accordingly.

6.3.1.2 Open and Closed Loop Power Control

The transmit powers of the BS and the MS must be dynamically changed because a variety of phenomena affect the quality of the signal. These include fading, velocity of the mobile, distance from the BS and so on. Also, there must be mechanisms by which the MS and BS can determine how the transmit power should be adjusted. This is achieved as follows.

Open loop power control is usually implemented on the reverse link. In open loop power control, the MS measures the quality of a reference channel from the base station. There may be a variety of measures such as RSS or frame or BER. If the RSS or BERs are above certain thresholds, the mobile terminal will automatically reduce its transmit power. If the signal quality is not good, the MS will increase it's transmit power. Clearly this is not a good mechanism for a variety of reasons. First, the decision on reducing or increasing the transmit power on the reverse channel is based on the measurement of the signal quality on the forward channel. These channels are not usually correlated, and a good signal reception on the forward channel does not necessarily mean the same on the reverse channel. The MS has no means of determining whether it has achieved the goal of minimizing the transmit power. Also, depending on the system, there may be significant delay in implementing this power control. In TDMA systems, the MS reception and transmission times are different, and there will be a lag time in implementing open loop power control.

Closed loop power control eliminates the disadvantages of open-loop power control by implementing a feedback mechanism between the BS and the MS. The BS measures the quality of the signal received from the MS and indicates what actions the MS should take via control signaling on the forward channel. Closed loop power control can also be used to control the transmit power at the BS. This is usually less important because the BS is not limited by battery power. However, adjusting the BS transmit power can benefit the system in terms of reducing the overall cochannel interference.

Example 6.9: Open Loop Power Control in IS-95 CDMA

In IS-95 (CDMA) standard, power control is extremely important because of the near-far effect described in Chapter 5. Because all the voice channels occupy the same time and frequency slots, the received signals from multiple users must all have the same RSS for each one to be detected correctly. An MS that transmits at a large power unnecessarily will essentially jam the signals of all the other MSs.

The open-loop power control scheme in IS-95 operates on the following principle. A mobile terminal that is closer to a BS should transmit less power because its signal suffers a smaller path-loss. Mobile terminals that are in deep fade or far away from a BS should transmit at a larger transmit power to overcome the loss in signal strength. Upon powering up, the MS adjusts its transmit power based on the *total* received power from all BSs on the forward channels [GAR00]. The BSs are not involved in the power control mechanism. The reference channel on the forward path used to determine the transmit power of the MS is the *pilot channel.* If the pilot channel is received very strongly, the MS transmits a weak signal to the BS. Else it transmits a strong signal to the BS.

Example 6.10: Closed Loop Power Control in GSM

The closed loop power control on the reverse link of GSM operates as follows [GAR99]. The MS measures the RSS and the signal quality of up to six neighboring BSs and reports its measurements to the base transceiver subsystem (BTS). The BTS also measures the RSS, signal quality, and distance to each of the mobile terminals in its serving area. From these measured values, it determines the minimum required transmit power and informs the MS of this value via a five-bit field in the slow associated control channel. The power control is performed in steps of 2 dB.

6.3.1.3 Centralized and Distributed Power Control

The open loop and closed loop power control mechanisms discussed earlier try to dynamically adjust the *transmit power* of the MSs or BSs based on the thresholds for SIR or BER set in the network. The goal of any power control scheme should be to uniformly render the SIR of all users to a value, which is usually the *maximum possible SIR* in the system. In terms of how such an optimization can be done are two approaches—centralized and distributed.

In a centralized power control (CPC) scheme, a central controller in the BS controller or mobile switching center has knowledge of *all* the radio links in the system. That is, the transmit powers, received powers, SIRs, and BERs for all mobile

terminal–BS combinations are known to this centralized controller. Assuming that the system is interference limited, an optimization algorithm can be implemented to maximize the minimum SIR in the system and minimize the maximum SIR in the system, thereby equalizing the SIR of all radio links [GRA93]. Although this provides an optimum solution, this scheme is extremely hard to implement because the centralized controller has to dynamically keep track of all the links in the system and compute the transmit powers for each mobile terminal.

In distributed power control (DPC) [GRA94], the mobile terminals adjust their transmit powers in discrete steps. This is similar to what is actually done in practice. For theoretical simplicity, it is often assumed that the MSs adjust their transmit powers synchronously. The power adjustments made by the MSs result in the transmit powers iteratively converging to the optimum power control solution. Ideally this should result in all the mobile terminals having the same SIR (which should be the maximum possible in the system) as in the CPC scheme after a number of adjustments. In practice, the adjustment to the power levels is also discrete (in steps of a few dB).

Example 6.11: Power Adjustment Levels in Example Wireless Networks

In GSM, each mobile terminal is required to either increase or decrease its power level by 2 dB depending on the message sent by the BTS. In CDMA, mobile terminals can change their power levels in steps of 1 dB. The IS-136 standard requires that a mobile terminal be able to change its transmit power by 4 dB in response to a command from the BS in 20 ms. In CDPD, the transmit power is set based on the signal power received and *not* the SIR. Consequently, power control is not based on signal quality. It is based only on the absolute signal strength.

6.3.2 Power Saving Mechanisms in Wireless Networks

In addition to dynamically changing the transmit power, there are several other mechanisms built into the operation of most wireless networks for saving the battery power of the MS. A variety of measurements have indicated that the most battery power is consumed during transmission of signals. A significant amount of power is consumed during active reception of signals (although this is lesser than the power consumed during transmission). A third mode of operation, called *standby* mode, consumes nearly an order of magnitude less power than either during transmission or reception of signals [AGR98].

Example 6.12: Power Consumption in Lucent WaveLAN

Lucent's WaveLAN (now Orinoco) is a WLAN product based on the IEEE 802.11 standard. It operates in the 2.4 GHz ISM bands. The power consumed by a 15 dBm WaveLAN radio is 1.825 W in the transmit mode, 1.8 W in the receive mode, and 0.18 W in the standby mode [AGR98]. Clearly, the standby mode operates with very little power and operating in this mode can save the battery life.

The operation of wireless networks is often designed to ensure that the mobile terminal spends as much time as possible in either a standby or sleep mode in order to conserve power. Several techniques are employed in wireless networks to reduce the amount of time spent in transmitting or receiving signals. In the case of laptops and other data terminals, a *sleep* mode of operation is preferred where the radio transceiver is shut off to conserve power. In voice networks such as GSM and IS-95, the *voice activity* factor is used to reduce either the transmit power or completely stop transmission when there is no speech activity. We discuss these techniques in the following sections.

6.3.2.1 Discontinuous Transmission and Repetition at Lower Transmit Powers

A particularly attractive option for saving the battery power of a mobile terminal is not to send information unnecessarily. With real-time communications, it is often assumed that there is a constant stream of data to be transmitted and also such communications are sensitive to delay and jitter. Usually, this is not the case, especially in a two-way conversation where one of the users is listening for some duration of time. Data communications are less affected by delay and jitter; it is possible to buffer data and transmit it at a later time. Discontinuous transmission is mostly employed in cellular telephone networks where additional hardware and algorithms are used to detect the presence or absence of voice. In ancient mobile telephones, some amount of information would be transmitted all the time, whether the person was actually speaking or not. With the use of *voice activity detection* (VAD), it is now possible for an MS to behave differently when no voice activity is detected. One of the possibilities (assuming ideal voice activity detection) is to not transmit any signal when the user is not speaking. A second alternative is to repeat data, but at a far lower signal power than usual. This will ensure that data is transmitted all the time, but the total power consumed corresponds to only that data which was actually generated by speech. VAD has its associated problems. In high-noise situations, the mobile terminal must be able to distinguish between the presence of useful signals in high noise and simply noise. Also it should be able to detect low-level voice activity. If VAD is not implemented correctly, there may be clipping of speech or an additional hangover after a talk spurt. Also, if there is absolutely no transmission, subjective tests indicate that silent gaps are extremely annoying. Consequently, systems usually insert a very low power *comfort noise* signal during silent gaps.

Example 6.13: Discontinuous Transmission in GSM

GSM transmits a comfort noise signal when there is no speech activity. When a VAD determines that there is no speech activity, the MS enters a hangover state to prevent clipping of speech due to very short silence periods. If there is no speech activity after the hangover period has elapsed, a silence identifier frame is transmitted at larger intervals than voice frames. The receiver will insert comfort noise when it detects the presence of the silence identifier frame.

Example 6.14: Discontinuous Transmission in IS-95

In IS-95, speech coders operate at different rates depending on the voice activity. The number of bits generated per frame will be different depending on the rate of the speech coder. The data stream corresponds to 9,600, 4,800, 2,400 or 1,200 bps. If traffic is generated at 9,600 bps, bits are transmitted at 100 percent of the transmit power. At lower data rates, the bits are repeated and then transmitted at half, one-fourth, or one-eighth of the transmit power on the forward channel. On the reverse channel, discontinuous transmission is employed.

6.3.2.2 Sleep Modes

A common approach for saving battery power is to allow the MS to enter into a *sleep mode* during periods of inactivity. The idea here is based on earlier discussion where we mentioned that the most power is consumed during signal transmission and significant power is consumed during signal reception as well. By entirely shutting off the RF hardware, it is possible to further reduce the battery consumption. However, there are problems associated with shutting off the RF hardware completely. What happens if a call or a packet arrives for a mobile terminal when it is shut off? The network should be able to make provisions for handling calls or packets that arrive for an MS that is in sleep mode.

Example 6.15: Sleep Mode in IS-136

In IS-136, the *standby time* is defined as the time for which an MS can be powered on and is available for service on a control channel before it needs recharging. The operation has been designed to allow the MS to enter a *sleep mode* for long periods of time when it is on standby. An MS is required to monitor the forward link only for a few time slots in order to determine whether there is a call addressed to it. The network may, however, require the mobile to monitor channels more frequently. The MS also has to monitor neighboring channels for handoff and monitor broadcast information. These will affect the time for which it can enter the sleep mode. For the rest of the time slots, the MS can enter a sleep mode.

Example 6.16: Sleep Mode in IEEE 802.11

In IEEE 802.11 WLANs, an MS can enter the sleep mode and inform an AP of its decision. Because this is a LAN and handoffs are less frequent than in cellular systems, handoff is less of a problem. Because of the bursty nature of data traffic, the arrival of packets addressed to the MS is a bigger concern. The AP buffers packets addressed to the sleeping mobile in 802.11. A beacon signal is transmitted periodically that contains information about buffered packets intended for sleeping mobiles. The MS wakes up at times when it expects the beacon and determines whether it should reenter the sleep mode or awaken completely to receive packets.

6.3.3 Energy Efficient Designs

The most common technique for conserving energy in a mobile terminal is to use advanced hardware design. Low power digital CMOS, mobile CPU microprocessors, and other hardware design approaches that consume very little power are usually employed in laptops and handheld computers. Beyond the actual hardware design, there are approaches at other layers that can enable savings on power consumption, thereby improving the lifetime of a battery. There are three approaches for improving battery life. The first approach is to tune the protocols employed in a wireless network to reduce power consumption. The second approach is to investigate power efficient modulation and coding techniques (see Chapter 3 for more details on modulation and coding). The final approach is in software design for mobile terminals.

6.3.3.1 Energy Efficient Protocols

Most protocols in data and voice networks have some relationship with the OSI protocol model even though they do not exactly match the seven layers. In wireless networks, the more important layers are the physical layer that handles the actual transmission of symbols and the link-layer that handles transmission of data packets or voice packets in link-layer frames and also controls access to the wireless medium. With the emergence of TCP and IP as the most popular transport and network layer protocols, these two are usually seen in wireless networks, although they are more restricted to data networks currently. As VoIP becomes popular, IP will make its presence in voice networks as well. As discussed earlier, the important power conservation principle in mobile terminals is to *minimize* the amount of time spent in transmitting signals. This principle can be applied to the design of different protocols in different layers.

6.3.3.2 Link Layer and MAC Design

At the link layer and in the design of medium access control techniques, power conservation factors should be taken into consideration. Two areas of design have been addressed in the literature. The first area is MAC design where the design goal is to eliminate unnecessary collisions and to employ better protocols for sleep mode and broadcast operation, as well as eliminating unnecessary processing at the MAC layer. The second area of design involves the link layer where ARQ error control schemes are employed for retransmission of lost or damaged packets.

Design at the MAC layer has focused on the following issues. Techniques to avoid retransmission due to collision as far as possible have been incorporated. Collision is not an issue in cellular telephony, where a channel is dedicated to the voice call for the duration of the call. However, in most data networks, such as IEEE 802.11, HIPERLAN, and CDPD, collision is an issue. Even in reservation-based schemes such as GPRS, the access to the network is achieved by a contention-based protocol.

Example 6.17: Collision Avoidance Mechanisms in Wireless Data Networks

Collision avoidance mechanisms have been incorporated in both IEEE 802.11 and HIPERLAN based on carrier sensing. In IEEE 802.11, there are two forms

of carrier sensing—at the physical layer and at the MAC layer. At the MAC layer, the length of transmission of a signal is detected via a field in the MAC frame and a *net allocation vector* (NAV) is set so that no signal transmission or physical carrier sensing is attempted in this period, thereby reducing the chance of a collision and also reducing the energy spent in monitoring the channel. In HIPERLAN, multiple contention phases eliminate the chance of collision to a great extent. In CDPD, the downlink carries a "status" flag that indicates whether the uplink is busy or idle. This once again reduces the possibility of collision.

Outside of collision avoidance, it is possible to use some intelligent techniques to further reduce unnecessary battery consumption. In WLANs, an MS will receive *all* packets regardless of whether they are addressed to it. If the packet is not addressed to the MS, it is discarded. This results in an unnecessary waste of battery resources. One possibility for improving this situation is to simply look at only the header information and continue receiving the signal only if the packet is addressed to the MS.

Example 6.18: Intelligent Processing in HIPERLAN

In HIPERLAN, some header information is transmitted via a low bit rate transmission scheme to reduce battery consumption [WOE98]. The reason for this is as follows. As the data rate increases, the effects of multipath delay spread require the use of equalization techniques as discussed in Chapter 3. Equalization schemes consume a lot of battery power. At 23 Mbps, which is the data rate supported by HIPERLAN, equalization becomes very important. In order to reduce battery power consumption, the header information is transmitted at a lower data rate (1.4706 Mbps). The entire header is not transmitted at a low data rate. Only a 34 bit value of the destination address is sent at this low data rate. The MS determines whether the received hash value matches its own hash value. If the hash value does not match, the rest of the packet is not received. If the hash value matches, the equalization circuitry is switched on, and the rest of the packet is decoded. Of course, there is a possibility that the packet still may not be intended for the MS, because the hash values are not unique. However, the chances of this happening are small. By not receiving signals or using the equalization circuitry unnecessarily, HIPERLAN terminals can save battery power significantly.

Other possibilities for intelligent packet reception are possible. Because the downlink is controlled in infrastructure networks by a BS or AP, this can schedule the broadcast of packets intended for different mobile terminals. The MS will then have to decode only those packets that arrive in the vicinity of its schedules reception time [AGR98].

ARQ schemes are employed at the link layer to retransmit data packets that are lost. ARQ schemes are not useful for real-time traffic such as voice. Packet losses can occur due to several reasons—collisions, interference, fading, and multipath delay spread. The collision avoidance mechanisms discussed earlier try to eliminate collisions to the extent possible. However, collisions cannot be entirely avoided as discussed in Chapter 4. Also, interference, fading, and other radio channel effects can result in errors in the received packet. Retransmission techniques are incorporated at the link layer based on error detection schemes. It is possible to

reduce retransmissions if error recovery can be performed at the receiver via forward error correction. In fact several wireless systems include *block interleaving* as discussed in Chapter 3 to reduce the effects of burst errors and enable forward error correction. However, if channels conditions are very bad, none of these techniques can recover from errors, and retransmission of packets will be necessary.

Retransmitting packets will not be useful if the channel conditions continue to be harsh. In fact this will result in unnecessary transmissions of packets that are bound to be lost or damaged. In [ZOR97], the energy efficiency of an error control protocol is defined as follows:

$$\lambda = \frac{\text{total amount of data delivered}}{\text{total energy consumed}} \tag{6.2}$$

Zorzi and Rao argue that this metric, which corresponds to the average number of packets delivered correctly during the lifetime of a battery, will influence the choice of ARQ protocols and that suboptimal protocols may in fact be better as far as energy consumption is concerned.

Example 6.19: An Adaptive Energy Efficient Go-Back-N ARQ Protocol

A classic Go-back-N ARQ scheme will transmit up to M packets and wait for acknowledgments from the receiver. The receiver will only accept packets in order and will send a negative acknowledgment (NAK) if a packet is not received. The receiver may also acknowledge several packets received correctly and in sequence with a single acknowledgment for the last correctly received packet. If the sender times out while waiting for an acknowledgment for packet N in the set of M packets or receives a NAK for the packet N, it will retransmit all the packets starting from N until M once again. All these packets may be lost if the channel conditions continue to be bad. Instead, the adaptive protocol suggested in [ZOR97] operates as follows. A probe packet is transmitted when there is a NAK or a timeout. The probe packet is a small packet that has minimum payload or simply a header so that the mobile terminal does not waste resources transmitting large packets. Only if a positive acknowledgment is received for the probe packet will the sender resume normal transmission of packets. Under slow fading conditions and small energy consumption for probing packets, this scheme can increase the energy efficiency by three times. Such a scheme can be worse than the regular scheme if the channel conditions are varying rapidly because the channel may have degraded immediately after an acknowledgment is received for the probe.

6.3.3.3 Transport Layer Design

The most common protocol employed at the transport layer in data networks is the TCP. As discussed in the previous section, as long as the channel conditions are bad, it is wasteful to transmit packets because they will not be delivered correctly. TCP has in-built mechanisms to back off when it detects packet losses. This backing-off of transmissions is initiated not because channel conditions are suspected to be bad, but because TCP assumes that there is congestion in the network, and transmissions should be reduced to ease congestion. However, it is possible that indirectly this may also aid in reducing unnecessary energy consumption in

wireless networks, especially when there are correlated errors on the wireless channel. In [ZOR99], the energy efficiency [defined in Eq. (6.2)] of TCP is investigated with this point of view. Analysis there indicates that depending on the nature of the channel, and the type of TCP implementation, TCP parameters can be tuned to increase energy efficiency significantly. In certain cases, it is possible to increase the energy efficiency by almost three times.

If parameters of TCP are not set correctly, the congestion avoidance mechanisms can degrade throughput and increase energy inefficiency. Split approaches for TCP [AGR98] introduce intermediate hosts in the network that keep track of missing packets and acknowledgments and handle TCP congestion avoidance mechanisms appropriately. These approaches can also increase the energy efficiency of the system.

6.3.4 Energy Efficient Software Approaches

Significant reduction of battery consumption can be obtained if the MS can be made to operate intelligently to reduce power consumption. Battery is consumed in mobile devices due to accessing of hard disks, operation of the CPU, and power consumption in the display in addition to the wireless communication unit and the communication protocols that we have considered so far. Energy management strategies are appropriate for each of these components, and this is usually achieved by the operating system (OS). These components usually have low power modes of operation and in many cases multiple modes of operation. A thorough discussion of these issues is provided in [LOR98]. The operating system will have to decide which mode of operation is appropriate at what time and when there should be a switch from one mode to another (*transition*). It should decide how a component's functionality can be modified to move it into low power modes as often as possible (*load change*) and also how software can be employed to permit novel power-saving use of such components (*adaptation*). Important factors in deciding strategies for any of these decisions involve what effect a strategy may have on the *overall* power savings because saving power in one component may affect the performance or power consumption of other components in addition to introducing unnecessary overhead. In certain cases, the lifetime of a battery will not be the sole issue, but how much of productivity can be obtained from a mobile terminal.

Table 6.1 provides a summary of results from [LOR98] that considers energy management issues related to secondary storage (hard disk, etc.), the processing unit, and the mobile terminal display units. A variety of power-saving strategies are possible as shown in the Table 6.1. The improvement in power saving varies depending on the type of mobile device, secondary storage device, processor, and display.

Significant power is consumed solely by access to hard disk, and it is suggested that it may be better to offload some of the storage onto a network file server. This assumes that the power consumed with the wireless transceiver will be smaller than the power consumed by disk access. Because transmission and reception does not involve mechanically moving parts, this may be a viable option, especially if energy efficient protocols are employed. Flash memory is nonvolatile

Table 6.1 Energy Management in Software in Mobile Terminals

Component	Secondary Storage	Processor	Display
Power-Saving Features	Five power modes—active, idle, standby, sleep and off	Clock slowdown, shut-off functional units, shut-off processor	Color to monochrome, reduce update frequency, turn display off
Issues	Sleep and standby modes consume far less power, but moving from standby to idle mode consumes power	Clock speed reduction can increase task time, thereby increasing power consumption; shutting off processor works best	Affects readability Can be annoying if there are flashes for updates
Transition Strategies	Enter sleep mode when inactive for some fixed threshold of time (several seconds) Use dynamic thresholds based on previous samples of disk access Predictive spin-ups of the disk (has not worked very well)	Process scheduler knows whether processes are ready to run or running. When processes are blocked, the CPU can be turned off (UNIX and Windows) Predict the number of busy CPU cycles and set the CPU clock speed	Turn off or turn down the display if there is no user input after a period of time
Load Change Strategies	Increase number of disk accesses—increase cache size Prefetch data based on prediction of usage before spinning disk down Reduce paging and improve memory access locality	Use lower power instructions (energy efficient compilers) Reduce the time taken by low-level tasks and pipeline tasks to units that can be turned off Reduce unnecessary tasks (block instead of busy-wait)	No formal load change strategies
Adaptation Strategies	Switch to flash memory for cache and storage Offload access to the network Reduce disk speeds in low power modes, improve energy consumption of disks	Design motherboards that automatically power down all components when the processor does not present them with a load	Use sensors to determine if a user is looking at a display or not Use lighter colors for display Provide ability to display only active windows and dim the rest of the display

and consumes less power. Against 0.9 W consumed only in idle mode for a hard disk, flash memory consumes 0.15–0.47 W for reading and writing. However, it is about 20 times more expensive than hard disks. Simulations indicate that flash memory used as secondary storage can reduce power consumption by 60 to 90 percent. Turning the processor off and reducing the voltage and clock speed of a processor can result in power savings that can be as large as 70 percent if the strategies are implemented correctly. The use of low-power states can further reduce power consumed by display units. Other components that can benefit from the OS employing low-power modes include sound cards, modems, and main memory [LOR98].

6.3.5 Implementation of Radio Resource and Power Management: A Protocol Stack Perspective

The purposes of the radio resources management (RRM) layer are threefold. First, it has to implement power control in the mobile terminals to reduce interference in the system. As a by-product this will increase the battery lifetime of the mobile terminal. Second, it has to assist the network and the mobile terminal to be connected on the best possible radio channel available to it within the cell. Last, it has to function to enable the mobile terminal and the network to handoff the mobile terminal's connection from one cell to another when the mobile moves across the boundary of a cell. All these tasks are performed by RRM entities that reside in the mobile terminal and sometimes in the network side of the wireless system.

The radio resource and power management functionality is usually handled by a management entity that interfaces across the lower three layers of the OSI protocol stack (as in the case of CDPD). Alternatively, it may be viewed as an application layer on top of the lower three layers (as in the case of GSM). In any case, handling of this functionality requires knowledge of exactly how signals are behaving at a particular point of time. This automatically requires feedback from the lower layers of the protocol stack.

In a manner similar to mobility management, there are two approaches to RRM. As we discuss in Chapter 9, in the case of GSM, the BS controller and the mobile switching center need to communicate with the mobile to obtain information about the state of the radio channel and provide the instructions for power control and selection of channels. A *bidirectional* logical channel is required for the communication between the mobile terminal and the network. In data networks, such as CDPD (discussed in Chapter 9), the mobile station has autonomous operation and decides for itself what should be the appropriate action. Power control is much harder to implement in this case. However, selection of radio channels and entering sleep mode of operation is done at the mobile station. In this case a *unidirectional* channel is sufficient. In CDPD, for example, the mobile–database station provides the mobile terminal information about available channels, transmit powers, and thresholds, but it does not expect any feedback from the mobile terminal in return. The mobile terminal is in charge of independently selecting channels and choosing what actions to perform. The protocol architecture and location of the RRM layers are described in subsequent chapters.

6.4 SECURITY IN WIRELESS NETWORKS

Wireless access to the Internet is becoming pervasive with diverse mobile devices being able to access the Internet in recent times. Current deployment is small, and the security risks are low as of today in these emerging technologies. However, the widely varying features and capabilities of wireless communication devices introduce several security concerns. The broadcast nature of wireless communications renders it very susceptible to malicious interception and wanton or unintentional interference. At least minimal security features are essential to prevent casual

hacking into wireless networks. Since the advent of analog telephony, wireless service providers have suffered several billion dollars of losses due to fraud. In this section, we address security issues in general. We provide an overview of network security services and mechanisms and describe some specific wireless examples.

6.4.1 Security Requirements for Wireless Networks

Security requirements for wireless communications are very similar to the wired counterparts but are treated differently because of the applications involved and potential for fraud. Different parts of the wireless network need security. Over the air security is usually associated with privacy of voice conversations. This is changing with the increasing use of wireless data services. Message authentication, identification, authorization, and so on also become issues here. Wireless networks are inherently insecure compared to their wired counterparts. The broadcast nature of the channel makes it easier to be tapped. Analog telephones are extremely easy to tap, and conversations can be eavesdropped using an RF scanner. Digital systems such as TDMA and CDMA are much harder to tap, and RF scanners cannot do the trick anymore, but as the circuitry and chips are freely available, it is not hard for someone to break into the system. Very little work has been done on optimizing security services for wireless systems, and patchwork solutions have made wireless networks not very secure. As long as the deployment is sparse and potential for harm to the consumer small, such measures can be sufficient. As more people use wireless access to the Internet and use wireless networks for e-commerce, credit-card transactions, and so on the potential for harm increases. In this section, we try to address some security requirements that have been identified for wireless voice networks and some that are emerging for wireless data networks.

6.4.1.1 Privacy Requirements of Wireless Networks

Privacy requirements are twofold in wireless networks. As discussed in Chapter 1, along with the air-interface, there is also a fixed infrastructure for handling the registration of mobiles, billing, mobility, power control, and other issues. There are privacy requirements for the air-interface and others for the messages transmitted over the wired infrastructure.

A variety of control information is transmitted over the air in addition to the actual voice or data. These include call setup information, user location, user ID (or telephone number) of both parties, and so on. These should all be kept secure because there is potential for misusing such information. Calling patterns (traffic analysis) can yield valuable information under certain circumstances. A flurry of calls between the CEOs of two major companies may indicate certain trends if it was discovered, even if the actual information in the calls was secure. Hiding such information is also important. In [WIL95], various levels of privacy are defined for voice communications.

At the bare minimum, it is desirable to have wireline equivalent privacy for all voice conversations. We commonly assume that all telephone conversations are secure. Although this is not true, it is possible to detect a tap on a wireline telephone. It is impossible to detect taps over a wireless link. To provide privacy that is equivalent to that of a wired telephone, for routine conversations it is sufficient to

employ some sort of an encryption that will take more than simple scanning and decoding to decrypt. Using DES with a 56-bit key would seem to be adequate for this purpose. In order to alert wireline callers about the insecure nature of a wireless call that is *not at all* encrypted, a "lack-of-privacy" indicator may be employed.

Wilkes [WIL95] calls these two levels of security as levels zero and one, respectively. Level-0 privacy is when there is no encryption employed over the air so that anyone can tap into the signal. Level-1 privacy provides privacy equivalent to that of a wireline telephone call, one possibility being encrypting the over-the-air signal. For commercial applications, a much stronger encryption scheme would be required that would keep the information safe for more than several years. Secret key algorithms with key sizes larger than 80 bits are appropriate for this purpose. This is referred to as Level-2 privacy. Encryption schemes that will keep the information secret for several hundreds of years are required for military communications and fall under Level-3 privacy.

For wireless data networks, a bare minimum level would be to keep the information secure for several years. The primary reason for this is that wireless electronic transactions are becoming common. Credit card information, dates of birth, social security numbers, email addresses, and so on can be misused (fraud) or abused (junk messages, for example). Consequently, such information should never be revealed easily. A Level-2 privacy will be absolutely essential for wireless data networks. In certain cases, a Level-3 privacy is required. Examples are wireless banking, stock trading, mass purchasing, and so on.

6.4.1.2 Other Security Requirements in Wireless Networks

Although privacy and confidentiality continue to be the important issue in wireless networks, other security requirements are becoming significant in recent times. There has been widespread fraud and impersonation of analog cellular telephones in the past. Although this is more difficult with digital systems, it is not impossible. There is thus a need to correctly *identify* and *authenticate* a mobile terminal. This becomes more important for private wireless data networks. For example, an organization that has installed a wireless LAN within its premises discovers that the coverage area extends into a neighboring street. With the reducing costs of wireless LAN PC-cards, anyone could buy a PC-card and access the organization's WLAN from the street. Privacy would assure that the person would not be able to read information being transmitted over the organization's communication network. However, the person could perhaps access the Web, thereby occupying the organization's bandwidth, or obtain an IP address from a DHCP server on the LAN if adequate authentication procedures are not employed. Similar situations exist with residential wireless networks, where it is important to keep security measures adequate, and yet not cumbersome to the layman.

6.4.1.3 Miscellaneous Issues

Even though traditional wired security measures are being put in place for wireless networks, wireless specific issues have been largely neglected. In addition to traditional security services such as privacy, authentication, message integrity, nonrepudiation, access-control, and availability, some of the wireless devices need certain

intermediate security services such as authorization, identification, and varying degrees or classes of security and privacy as discussed earlier. The potential security implications and interaction between security requirements and wireless network/device limitations are unclear. Wireless communication devices are expected to be mobile and have the additional requirement that they must consume as little power as possible while performing computations for encrypting or decrypting data to conserve battery power. This is a significant issue because cryptographic algorithms are computationally intensive and may drain the battery of a mobile terminal quickly. Because the spectrum is scarce, cryptographic protocols should also not waste resources by requiring several handshakes between the mobile terminal and the fixed network. This requirement is usually contrary to security services as they are implemented in wired networks. The wireless channel is error prone and may also result in messages being lost, duplicated, or damaged. It is also not clear what effects this may have on the overall performance of security protocols. Interference, fading, disconnections, handoff and other mobility-related procedures, and other peculiarities of the wireless network require robust security services that are at the same time resource efficient. These issues are yet to be addressed in wireless security.

6.4.2 An Overview of Network Security

Network security has made tremendous strides with the explosive growth of the Internet, and several textbooks have been written on this topic [STI95, STA98, SCH94]. It is beyond the scope of this section to present all relevant concepts in detail. The goal here is to present sufficient detail to enable the reader to appreciate and understand the subject material that follows in later sections.

6.4.2.1 Security Services

The primary method of understanding and designing solutions to a particular problem is to first develop a set of requirements that must be satisfied. In network security, classifying and defining various features that must be available to a network to keep information secure accomplishes this. Such various features are commonly referred to as security services. Although there are several ways of defining security services, usually we identify them as specific measures employing *security mechanisms* that combat security attacks on a network. Based on [STA98], we present a list of security services as follows.

Confidentiality or *privacy* is a security service that provides resistance to the security attack known as *interception*. This is the most intuitive form of security service where two communicating parties do not wish to reveal the contents of their transactions to a third party. In more rigid cases, the existence of the communication itself must not be revealed to unauthorized entities. Encrypting the messages and the identities of the two communicating parties is the most popular method of providing confidentiality. *Message authentication* is a term that is used for a security service that provides *integrity* of the message and also provides a guarantee that the sender is who he or she claims to be (*sender authentication*). The corresponding security attacks are *modification* of the message and *impersonation* of the sender's identity. Message authentication can be provided by attaching a digest of the message, which is encrypted by a key known only to the communicating parties.

Nonrepudiation corresponds to a security service against denial by either party of creating or acknowledging a message. This service is similar in nature to a signature by the creator of a document and *digital signatures* based on public key encryption schemes discussed below are employed to provide this service. The corresponding security attack in this case is *fabrication*. *Access control* enables only authorized entities to access resources. *Masquerading* is the security attack corresponding to this service. *Availability* ensures that resources or communications are not prevented from access or transmission by malicious entities. *Denial of service* is the attack corresponding to the security service of availability.

Although the security services discussed are the most prominent ones, other security services also play a role in certain applications. *Authorization* is sufficient in certain cases instead of a sender authentication. Authorization allows a user or communication command to execute certain operations. Such a security service would be sufficient, for example, to instruct a coffee machine to execute certain operations. *Identification* is another security service often used in transactions such as automatic teller machines. Such a security service would be useful in other applications such as credit card transactions over a WAP-enabled microbrowser.

6.4.2.2 Security Mechanisms

There are several ways in which security mechanisms can be provided. However, there is no single technique that can provide all required security services. *Encryption* is a technique that is employed both in wired and wireless networks for providing several security services. Encryption, when employed in clever ways, can provide confidentiality, message authentication, nonrepudiation, access control, and identification. Availability cannot be guaranteed by encryption because encryption is powerless against attacks such as cutting wires or jamming signals. In succeeding subsections, we shall see how encryption can be used to provide security services.

6.4.2.3 Confidentiality or Privacy

In order to keep a message confidential, the easiest technique is to scramble the message using a *key* so that if the message falls into wrong hands, the adversary will not be able to understand or descramble the message. The scrambling technique is usually called *encryption*. The message is referred to as *plaintext* or *cleartext*, and the encrypted version of it is referred to as *ciphertext*. It is common to denote two communicating parties as Alice and Bob and the adversary or opponent as Oscar. Mathematically, the encryption of a plaintext x into a ciphertext y using a key k is written as:

$$y = e_k(x) \tag{6.3}$$

The corresponding decryption is written as:

$$x = d_k(y) \tag{6.4}$$

Ideally, we would like the encryption scheme to be such that it cannot be broken at all. Because there are no practical methods of achieving such an unconditional security, encryption schemes are designed to be *computationally secure*. The

encryption scheme has to be powerful, in that given significant computational resources, an adversary must not be able to either find the key or decrypt the message in a reasonable time. Alternatively, if either the key or the plaintext can be determined in a short time, it should cost the adversary much more than what the value of the secret information would be to him. Usually, it is assumed that Oscar has knowledge of how the algorithm works, but not the key. Also, because of the standard formats of data packets and control messages in voice networks, Oscar usually has access to a limited number of plaintext-ciphertext pairs that he can use to perform a *known plaintext* attack to recover the key. Once the key is recovered, all subsequent ciphertext can be decrypted easily.

So far, we have discussed security services, and we have said that encryption can provide some of these services. What we have not discussed are the encryption algorithms that are employed or can be employed within these mechanisms. The details of these algorithms are beyond the scope of this book, and the subject matter forms a wide area of interest in itself. In this section, we briefly mention some of the algorithms that are in wide usage today. We also discuss the key sizes that are required to make these algorithms secure.

6.4.2.4 Secret-Key and Public-Key Algorithms

Encryption schemes have been available through the ages and have all been what are known as *secret-key* algorithms. Here, the communicating parties (Alice and Bob in Figure 6.13) share a secret key that they use to encrypt any communication between themselves. Usually, the encryption and decryption algorithms use

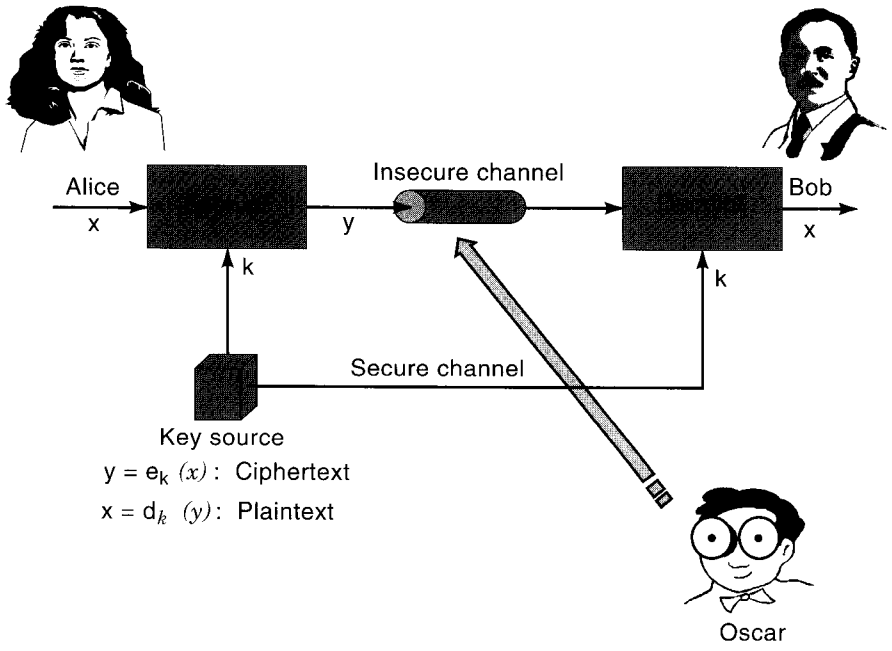


Figure 6.13 Conventional encryption model.

the *same* key, and, hence, such algorithms are also called *symmetric key* algorithms. Block ciphers such as the data encryption standard (DES) also fall under this category. Figure 6.13 illustrates a schematic of a conventional encryption scheme. The opponent Oscar has access to the insecure channel and thus the ciphertext. However, he has no knowledge of the secret key k shared by Alice and Bob.

Secret-key algorithms such as DES are based on two principles: confusion and diffusion. The former introduces a layer of scrambling that creates confusion as to what exactly might be the transmitted message. The latter creates a randomness whereby the effect of changing a small part of the plaintext message will result in changing half of the encrypted ciphertext. This eliminates matching patterns or frequencies of occurrence of messages. Most secret-key algorithms are thus unbreakable except by brute force [SIL99]. If the length of the key of a secret-key algorithm is n bits, at least 2^{n-1} steps are required to break the encryption. Today, a key length of 80 bits is considered to be sufficiently safe from brute force attacks even though a key size of 128 bits is usually recommended.

Example 6.20: Security of DES against Brute Force Attacks

DES is a block cipher that encrypts plaintext messages in blocks of 64 bits using keys that are 56 bits long. The total number of keys is 2^{56}. On average, half of them will have to be examined to determine the right key if a known plaintext-ciphertext pair is available. If a 500 MHz chip is employed for this attack, and one decryption (or encryption) can be performed in one clock cycle, to test 2^{55} keys, it will take $2^{55}/(500 \times 10^6)$ seconds = 834 days to break the encryption. This is not very secure if 834 chips are used in parallel, because the key can be obtained in a single day. The total cost will be about $16,680 if each chip costs $20!

Example 6.21: Security vs Advances in Chip Speeds

DES was broken in less than a day in January 1999 at a cost of $500,000. Today it is virtually impossible to break a well-designed block cipher with key sizes of more than 80 bits (which translates into examining around 280 keys by brute force). However, a common assumption (called Moore's Law) is that processor or chip speeds double every 18 months, thereby weakening any encryption scheme with time. For example, using a speed of 500 MHz for today, in 100 years, an encryption scheme that employs key sizes twice that of DES (i.e., 112 bits) can be broken in a day.

Example 6.22: Key Sizes in Wireless Systems

The key sizes used in current wireless systems are not sufficiently large enough for good security. In IEEE 802.11, a 40-bit key is used in the encryption algorithm (called wired-equivalent privacy or WEP) that is not secure by today's standards. IS-136 uses a 64-bit A-key that is more secure, but still considered weak.

The primary advantage of secret-key algorithms is that they are fast, and at the huge data rates that are being supported by today's networks, it is virtually

impossible to employ *public-key* algorithms. However, because every pair of users has to have a key, for a communication system with N users, at least $N(N-1)/2$ keys need to be created and distributed. This is not a trivial exercise and has its own weaknesses.

Example 6.23: Number of Keys with Symmetric Key Encryption Algorithms

Assume a small corporate network with 500 computers. A total of 124,750 keys are required (one between each pair of computers). Each computer needs to store 499 keys associated with the remaining computers. Suppose an employee gets a new handheld personal computing device. Not only will this handheld device need to load 500 new keys, but the remaining 500 old computers also need to be each updated with a key for the handheld computer.

There are techniques of key distribution for symmetric key algorithms such as the Needham-Schroeder key distribution scheme and Kerberos [STA98]. All these schemes need several handshaking steps and also an initial configuration of computers with *master keys*. Such master keys can be distributed physically in a secure manner. However, key distribution is still a potential weakness in the system. Another way of generating fresh keys for each communicating session is to use the master key and a one-time random number (called nonce) as inputs to a one-way hash function to generate a key. Alternatively, the nonce can be encrypted using the master key.

Example 6.24: Generation of Secret Keys in Wireless Networks

Most wireless networks (including cellular networks) and several wired networks employ identification schemes that depend on a *shared master key* followed by secure forms of hash algorithms to generate "fresh" keys that can then be used with a secret key algorithm to provide various security services. In general hashing a random number concatenated with a secret identifying parameter known only to the communicating parties can securely generate keys. In most cases (like IEEE 802.11), the size of the identification parameter (PIN number or master key) and the algorithms employing it provide loopholes and vulnerability in the protocol. It is generally suggested that the shared secret (whether a password or a PIN) should be at least as long as 80 to 128 bits (and the hash algorithm employed should have an output of at least 160 bits) because of a square-root attack called the "birthday attack" [STA98].

Block ciphers such as DES and the *advanced encryption standard* (AES) encrypt blocks of data at a time. Stream ciphers encrypt bits or bytes of data [STI95]. The advantage of stream ciphers is that there is no need for buffering data up to the block size and for padding. Stream ciphers may also be more suitable for a jitter sensitive voice conversations. The disadvantage is that these have to be used carefully because encryption with stream ciphers uses simple XOR operation.

Example 6.25: Use of Stream Ciphers in IEEE 802.11

In IEEE 802.11, the encryption algorithm RC-4 [STA98] is used to generate a pseudorandom key stream using a 40-bit master key and an initial vector (IV). The data are then simply XOR-ed with the key stream to create the ciphertext.

DES was the secret-key encryption standard for over 20 years. The National Institute for Standards and Technology (NIST) examined proposals for an AES in 1998. Of the five candidate algorithms, NIST selected Rijndael as the algorithm for AES in October 2000. AES is being considered for use in IEEE 802.11. A variety of factors were considered by NIST to determine the suitability of the algorithm for a standard. Security—resistance to cryptanalysis, mathematical soundness, randomness of the algorithm output; cost—licensing requirements, computational efficiency on various platforms, memory requirements; and algorithm implementation characteristics—ability to handle variable key sizes and block lengths, implementation as stream ciphers and hash functions, hardware and software implementations, and algorithm simplicity are three categories used for evaluation of these algorithms.

In addition to these standards, several freeware and other secret-key algorithms are available such as IDEA, RC-4, and Blowfish [STA98]. RC-4 in particular has been widely employed in Web browsers, as well as in wireless networks such as IEEE 802.11. Many hash algorithms and MAC algorithms exist that are based on secret-key encryption schemes. The secure hash algorithm (SHA) and the hashed MAC (HMAC) are used widely over the Internet for message authentication. The encryption algorithms employed in GSM and the North American digital cellular standards are proprietary.

Public key encryption is a radical shift in the way data is encrypted. Diffie and Hellman introduced the concept in 1977. With secret key algorithms, we have a situation that is similar to having a locked mailbox for *each pair of users*. Both users associated with a mailbox share a key that can unlock or lock the mailbox. Consider Figure 6.14. Here, if Alice desires to communicate with Dan, she unlocks the mailbox shared between her and Dan, deposits the message, and locks the mailbox again. The message is now accessible only to Alice and Dan who also has an identical key.

Clearly the number of mailboxes required for N users like Alice and Dan is $N(N-1)/2$. For example, we have six mailboxes for four users as shown in Figure 6.14.

The situation described is not the natural way in which we employ mailboxes. Mailboxes are *associated with individuals* and not pairs of communicating parties. The natural way to employ a mailbox is as described in the following example. Alice owns a mailbox. Only she has a key to lock or unlock the mailbox (i.e., only Alice has complete control over the mailbox). *Any other person* who wishes to communicate with Alice will deposit the message through a *slot* in the mailbox. Once the message is deposited in the slot, *only Alice has access to it.* Even the originator of the message cannot retrieve it, although he or she may regenerate the message from knowledge of the contents.

Public-key algorithms are similar to this example. Each individual has a pair of keys—the public key and the private key. As the name suggests, everyone knows the public key. So anyone can employ the public key to encrypt a message intended for the owner of the key. The public key is like the slot in the mailbox. Only the owner knows the private key. As a result, once the message is encrypted using the public key of the owner, only the owner can decrypt the message. Not even the originator of the message can decrypt it once the message has been encrypted.

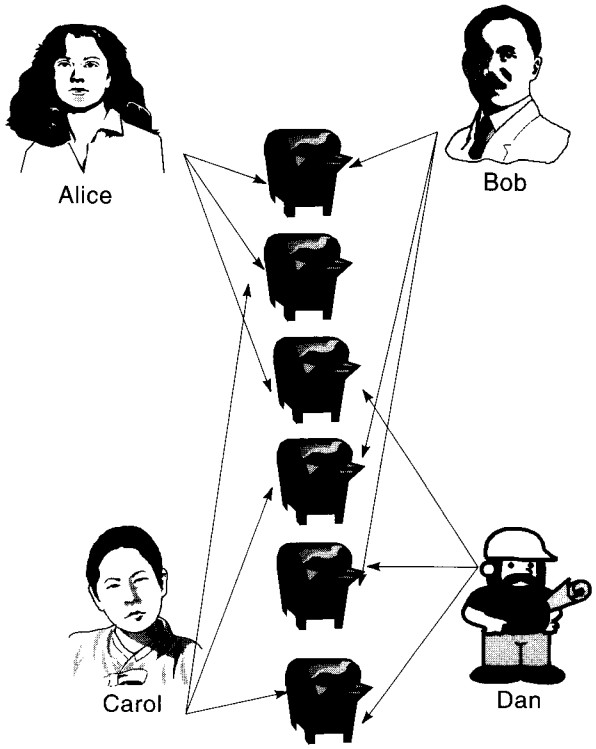


Figure 6.14 Multiple mailboxes with secret-key encryption.

Figure 6.15 shows the schematic of a public-key encryption scheme. Note that there is no need for secure transfer of the key anymore. Alice encrypts a message intended for Bob with Bob's public key $K_{pub,bob}$. The ciphertext is decrypted by Bob via his private key. The design criterion for public-key algorithms is as follows. Given a function $f(k,x)$, the following properties always hold.

- It is extremely easy to compute $y = f(kpub,x)$.
- Given *kpub,* and *y,* it is computationally not feasible to determine $x = f^{-1}(kpub,y)$.
- With a knowledge of *kprv* that is related to *kpub,* it is easy to determine $x = f^{1}(kprv,y)$.

Example 6.26: Trapdoor One-Way Functions

Functions that have these properties are called trapdoor one-way functions. Examples are the factorization problem and the discrete logarithm (DL) problem. The former is based on the fact that it is easy to multiply prime factors to arrive at a composite number (e.g., it is easy to find $7 \times 17 \times 109 \times 151 = 195{,}821$, but it is quite a hard task to split 30,616,693 into its prime number factors). The latter is based on the fact that it is easy to determine what 2^{23} mod 109 is (the answer

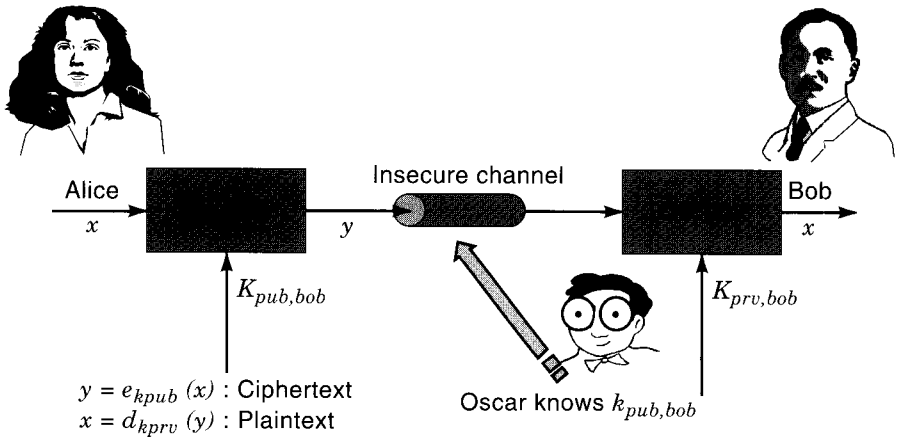


Figure 6.15 Public-key encryption scheme.

is 77). It is quite hard to find out what u is given $2^u \bmod 109 = 68$. Note that with real number arithmetic, it would have been trivial to determine u as $u = \log_2 68$. The modulo function which reduces the operations to be one set of numbers that are nonnegative integers less than 109 makes this problem very hard to solve. Integer factorization is employed in RSA, and the DL problem is used in the Diffie-Hellman Key Exchange Protocol and Digital Signatures.

RSA has been the most popular public-key algorithm. It employs integer factorization. The Diffie-Hellman key exchange protocol based on discrete logarithms is also very popular in wireless networks. This protocol is described in Appendix 6A and is commonly employed for key exchange for Web transactions, e-commerce, and IP security. The digital signature standard (DSS) is also based on discrete logarithms. Signature schemes based on RSA are also widely employed.

However, in the case of public-key algorithms, Oscar, the opponent, is aware of Bob's public key, and this adds an additional parameter to the problem. Because public-key algorithms are based on mathematical structures, for small key sizes, there are well-known results or tables that can be employed to break the encryption. As such, the key sizes are extremely large compared with secret-key algorithms. Today, for good security, public-key algorithms need keys that are three to 15 times larger than their secret-key counterparts. Because the mathematical bases on which public-key algorithms work are well known, they are susceptible to analytical attacks and require much larger key sizes compared with secret key algorithms. The mathematics of elliptic curves are also being employed in encryption schemes because they need smaller key lengths compared with RSA. Table 6.2 presents the key lengths and the time required to break some of the well-known public and secret-key algorithms [SIL99]. The values in this table are based on the assumption that $10 million is available for computer hardware. The key sizes in each row are equivalent.

The mathematical operations for public-key algorithms are also quite computationally intensive. Consequently, the encryption rates are quite small and public-

Table 6.2 Cost Equivalent Key Lengths (in Bits) of Various Encryption Schemes

Secret-key Algorithm	Elliptic Curve	RSA	Time to Break	Memory
56	112	430	Less than 5 minutes	Trivial
80	160	760	600 months	4 Gb
96	192	1,020	3 million years	170 Gb
128	256	1,620	10^{16} years	120 Tb

key algorithms are rarely used for bulk data transfer. Instead, they are employed to exchange a *session key* between a pair of communicating entities who will then use the session key with secret-key algorithms for the duration of that communication alone. This ensures that a new session key is employed each time a communication is initiated, thereby reducing the possibility of an adversary breaking the encryption scheme.

6.4.2.5 Message Authentication

Message authentication is a security service that provides two functions: sender authentication and message integrity. By sender authentication, what we mean is that the receiver can be assured that the message has been originated from the person who claims to have sent the message. Message integrity assures a receiver that no one has modified the message in transit. Both these functions can be accomplished by adding a *message digest* (MD) or MAC to a message. The MAC here should not be confused with the medium access control layer discussed in Chapter 4.

Example 6.27: Message Authentication in IEEE 802.11

In IEEE 802.11, as we saw before, the packet is encrypted using a stream cipher. If the key stream is not correct, upon decryption, the CRC (cyclic redundancy check) in the packet will fail and the access point can discard the packet. However, this implementation can be easily broken. Additionally, it is possible to filter packets based on the 48-bit 802.11 MAC address. Once again, it is not too hard to spoof the physical address of the device.

Secret-key algorithms can be employed for this purpose. The way a MAC is used to provide message authentication is as follows. It creates a fixed length sequence of bits that depend on the message itself and a secret key shared between the communicating parties. Irrespective of whether the message is a few kilobytes long or hundreds of megabytes, the MAC creates a sequence of bits of fixed length that directly depends on the message and the key. This sequence of bits is appended to the message, and then the result is transmitted over the insecure channel. Note that the message could be sent in plaintext form if confidentiality is not an issue. It is computationally infeasible to create a replica of the MAC without the message and key. If the message is modified in transit, the receiver can discover this fact by creating a MAC from the received message and comparing it with the

transmitted MAC. Because the secret key is shared only between the communicating parties, it also assures the receiver of the origin of the sender.

An MD operates in a slightly different manner. The MD depends only on the message and not the key. Hash functions are used to create message digests. The message is appended with the MD, and the result is encrypted using a session key shared between the communicating parties. This way, both the message and the MD, which verifies it, are kept secure. The MD has to be sufficiently long to prevent what is known as the "birthday attack." Given a message digest of length b bits, with a good probability, a fake message with the same MD can be generated in $2^{b/2}$ trials. This result is due to the fact that good probabilities of finding two people with the same date of birth exist in a group with roughly the square root of the number of days in a year. That is, in a group of 20 people, it is quite likely that any two would have the same birthday.

Figure 6.16 shows a schematic of message authentication with hash functions. On the left-hand side, Alice concatenates the message x and its hash value $h(x)$ together before encryption the result with the secret key k. The ciphertext $y = e_k(x||h(x))$ is transmitted over an insecure channel. Bob decrypts the ciphertext y and expects to find a message and its hash value concatenated together. He separates the message x from the hash value, computes a new hash value, and compares the two together. If the ciphertext is modified or replaced in between, Bob is able to discover this fact easily. No one can impersonate Alice because it is computationally impossible to create a ciphertext that decrypts into a message and its hash value without knowledge of the key k. Thus both sender authentication and message integrity is assured. The interested reader is referred to [STA98] for other schemes for message authentication. Using the hash function is generally preferred because of its speed.

6.4.3 Identification Schemes

Encryption schemes and hash functions are widely employed in password protection schemes and access control lists that are used for access control—the ability to allow or deny people access to certain resources based on their identification.

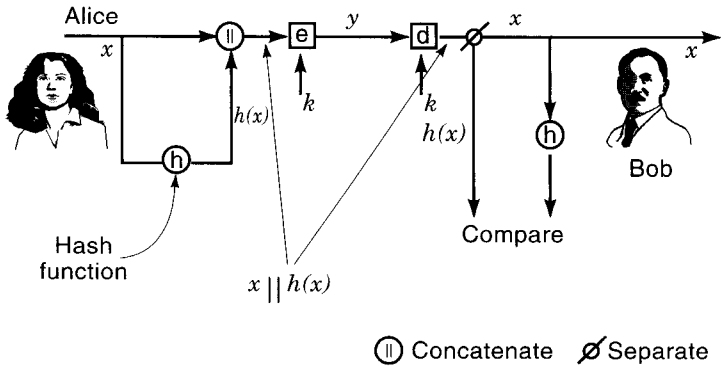


Figure 6.16 Message authentication with hash functions.

Identification by itself is an important security service that needs to be provided for a variety of applications. Access to an automatic teller machine, logging on to a computer, identifying the user of a cellular telephone to the network, etc. involve identification schemes. Note that there is a difference between identification and authentication. When we talk about authentication, there is usually some information-containing message that is exchanged between the parties and one or both parties need to be authenticated. Identification schemes (sometimes referred to as *entity authentication*) involve real-time verification of a party's identity and *need not* involve exchanging information-bearing messages.

Weak identification schemes are based on passwords or pin numbers that are time invariant. Usually the password or pin value is compared with a securely stored hash value. Such schemes are easily susceptible to replay attacks, especially if the password or pin is transmitted over the air in an insecure manner. *Challenge-response* identification or *strong identification* schemes are usually employed in wireless networks. Here, Alice proves her identity to Bob by demonstrating knowledge of a secret, rather than presenting the secret itself. For this purpose, a quantity called the "nonce" is used. A nonce is a value employed no more than once for the same purpose and eliminates replay attacks. Random numbers, time stamps, sequence numbers, and so on are used as nonce in practice. One example of a challenge-response protocol is as follows:

1. Alice is registered with Bob via a password and user name.
2. Bob sends Alice a random number (challenge).
3. Alice replies with an encrypted value of the random number where the encryption is done by using her password as the key (response).
4. Bob verifies that Alice indeed possesses the key (the password).

An eavesdropper Oscar cannot replay the response because the challenge is different if he tries to contact Bob. Oscar also cannot determine the password because the encryption scheme is sufficiently strong and the password is *never* revealed.

Example 6.28: Challenge Response Schemes in Cellular Networks

Figure 6.17 shows the architecture of the IS-136 digital TDMA standard. Details of this architecture are discussed in Chapter 8. The lower half of the figure shows part of the challenge response mechanism implemented in the IS-41 standard for network operations in IS-136. The network (Bob) generates a random number RANDU and sends it over the air to the mobile terminal (Alice). The mobile terminal computes a value AUTHU using the encryption algorithm called CAVE (Cellular Authentication and Voice Encryption) algorithm. The value AUTHU is transmitted over the air. The network computes its version of AUTHU and compares the two values. If the values match, the mobile terminal is identified (authenticated in the terminology of IS-41).

In order to look at the security mechanisms in specific technologies, details of the architectures of these technologies need to be described. We consider such details in subsequent chapters.

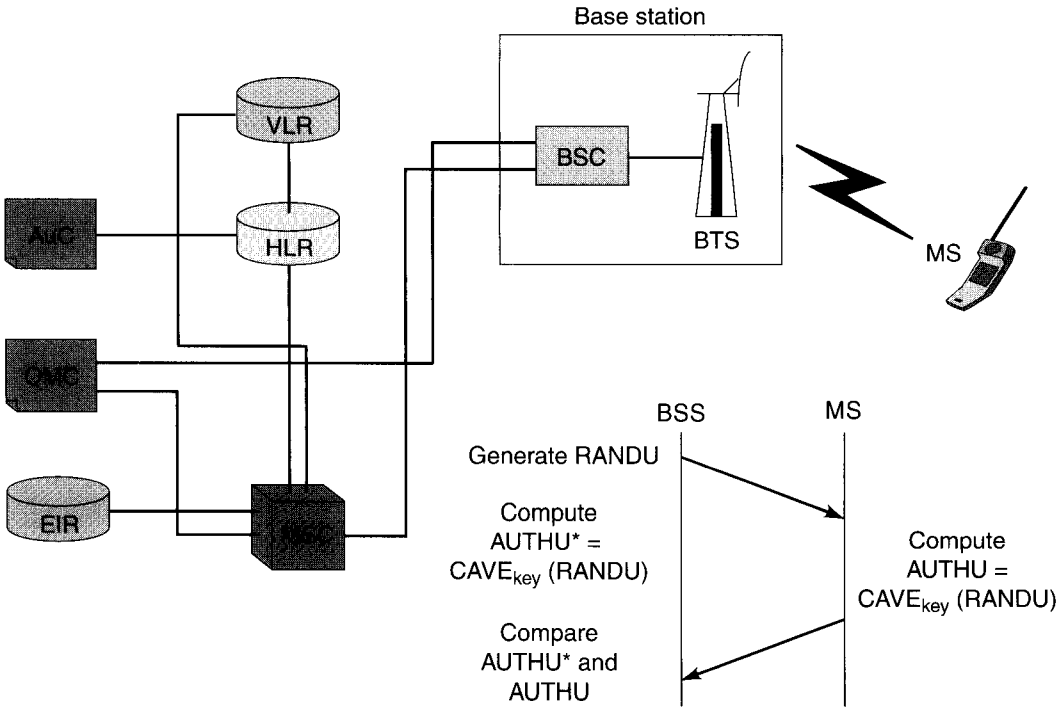


Figure 6.17 Challenge-response mechanism in IS-41.

APPENDIX 6A THE DIFFIE-HELLMAN (DH) KEY EXCHANGE PROTOCOL

The DH key exchange protocol is based on the DL problem discussed in Example 6.4. Let us suppose that Alice wishes to exchange a session key with Bob without sharing any secret with him. Alice chooses a base α and a large prime number p that are publicly known. She only chooses a random private number a. She computes $k_{pubA} = \alpha^a \bmod p$ which she sends to Bob. Note that given k_{pubA}, α, and p, it is computationally impossible to determine a. Similarly, Bob chooses a private random number b and computes $k_{pubB} = \alpha^b \bmod p$ which he transmits to Alice. Once again, it is extremely difficult to determine b. After obtaining the public keys of each other, Alice and Bob raise these public keys to the exponent corresponding to their private numbers respectively. That is, Alice will compute

$$k_s = k_{pubB}{}^a \bmod p = \alpha^{ab} \bmod p \tag{6A.1}$$

Bob computes

$$k_s = k_{pubA}{}^b \bmod p = \alpha^{ab} \bmod p \tag{6A.2}$$

This way both Alice and Bob have generated a common session key. An adversary Oscar cannot determine this key without solving the DL problem. At least, there

is no known solution for obtaining the session key other than by solving the DL problem.

APPENDIX 6B NONREPUDIATION AND DIGITAL SIGNATURES

We considered sender authentication and message integrity in this chapter. This does not, however, assure nonrepudiation. For instance, let us suppose that Alice is a consumer and Bob an e-commerce service provider. Bob claims that Alice placed an order with him for purchasing books worth $350, and Alice denies the transaction. Alice claims that she had requested books only worth $100. Both of them are able to produce ciphertexts and messages purportedly used in the transaction. Because both parties know the shared session key, it is impossible to verify who is being truthful and who is not. Public-key algorithms and digital signatures can be employed to resolve such situations.

We know that *only* the owner of the key knows the private key part of a public key algorithm. Consequently, this information can be used to bind the owner to a message transmitted by him. Popular public-key algorithms operate such that it is possible to encrypt a message using a private-key as well. We can compare this with the following scenario. Only the owner of a mailbox can slip a message through the slot because only he has access to the private key that opens the mailbox. No one other than Alice can encrypt the message using her private key (or produce a meaningful ciphertext that can be decrypted with her public key). The problem with this encryption is that *anyone* will be able to decrypt the message because the public key dual is available to everyone. But this is exactly the concept of a signature. If Alice were to sign a document, this means that anyone should be able to verify her signature. However, no one should be able to forge her signature. This is indeed the case here when a message is encrypted using the private key.

Digital signatures take the concept a step further. The entire document *need* not be encrypted. As already discussed, this process would be extremely slow. Instead a message digest of the message is "signed" or encrypted with the private key. The encrypted "signature" is appended to the message. Once again, because it

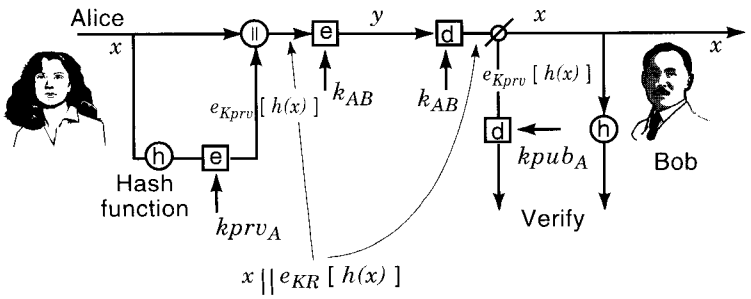


Figure 6B.1 Digital signatures

is computationally impossible to derive a message from the hash, the signature and the message are bound together. If the document needs to be confidential, the usual encryption procedures can be employed after the signature is applied. Figure 6B.1 shows a schematic of how digital signatures are applied. Here, k_{AB} is a session key that is used to keep the document confidential.

QUESTIONS

6.1 What three operational issues are important in wireless networks compared with wired networks? Why?
6.2 Name the two important issues in mobility management.
6.3 What is location management? What are the three components of location management? What are the tradeoffs between them?
6.4 Name three location update mechanisms.
6.5 Name three paging mechanisms. Explain what blanket paging is.
6.6 What are the two steps in handoff?
6.7 Explain three traditional handoff techniques.
6.8 Differentiate between mobile-controlled and mobile-assisted handoff.
6.9 Explain a general handoff procedure. Explain the entities involved with GSM as an example.
6.10 What is agent advertisement? Why is it important in Mobile IP?
6.11 What are smooth handoffs? What application(s) may benefit from them?
6.12 Why is power control important in wireless networks?
6.13 What are the differences in power control for voice-oriented and data-oriented networks?
6.14 Differentiate between open-loop and closed-loop power control.
6.15 Differentiate between centralized and distributed power control.
6.16 Name two types of power-saving mechanisms.
6.17 Differentiate between sleep modes in IS-136 and IEEE 802.11.
6.18 What intelligent protocol features are available in IEEE 802.11 and HIPERLAN to save battery power?
6.19 Describe some energy efficient software approaches.
6.20 What are the privacy and authentication requirements of wireless networks?
6.21 How are public-key and secret-key algorithms different?
6.22 Explain the importance of key sizes in the security of an encryption algorithm.
6.23 What is a challenge-response scheme? How does it work in IS-136?

PROBLEMS

6.1 A mobile terminal samples signals from four BSs as a function of time. The times and signal strengths (in dBm) from the samples are given in Table 6.3. Assume the mobile terminal is initially attached to BS 1 (BS_1). The mobile makes handoff decisions by considering the signals from the BSs after each sampling time. For example, if just RSS is used, just after t = 12.5s, the mobile terminal would be connected to BS_3. On a plot,

show the handoff transitions between BSs for each of the following algorithms as a function of time. If a condition is met for more than one BS, assume the best one (strongest RSS) is selected.

a. Received signal strength (RSS)

b. RSS + threshold of -60 dBm

c. RSS + hysteresis of 10 dB

d. RSS + hysteresis of 5 dB + threshold of –55 dBm

Table 6.3 RSS from Four Base Stations

Time(s)	0	2.5	5	7.5	10	12.5	15	17.5	20
BS_1	−47	−57	−52	−55	−60	−62	−60	−65	−64
BS_2	−59	−56	−55	−54	−52	−51	−49	−60.5	−52
BS_3	−70	−72	−75	−70	−58	−50	−60.5	−62	−75
BS_4	−72	−71	−65	−60	−55	−53	−50	−49	−56

6.2 In Problem 6.1, which technique is the best in terms of reducing the number of unnecessary handoffs? What parameters will you change to reduce the number of unnecessary handoffs? If the minimum required RSS for good signal quality is −55 dBm, would your answers change?

6.3 A mobile node has a home address of 136.142.117.21 and a care-of address of 130.216.16.5. It listens to agent advertisements periodically.

a. The agent advertisement indicates that the care-of address is 130.216.45.3. What happens? Why?

b. The agent advertisement indicates that the care-of address is 136.142.117.1. What happens? Why?

6.4 Mobile terminals in seven co-channel cells (labeled A) of a cellular system are transmitting on the same frequency channel as shown in Figure 6.18. Due to a glitch in the handoff mechanism, the transmit power of the mobile in the center cell increases without control as it moves away from a BS, and it continues to be connected to it even after

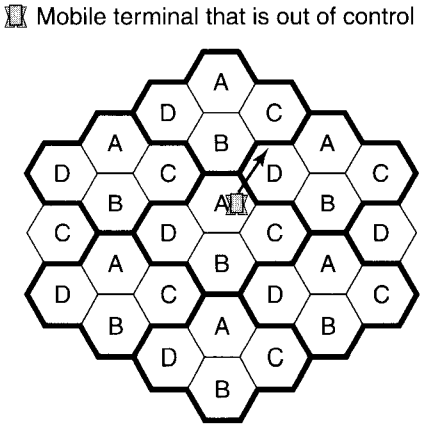


Figure 6.18 Mobile terminals of a cellular system.

it moves out of cell A into cell D as shown in the figure so that its transmit power is now three times as the rest. Assume that the frequency reuse factor is $N = 4$. Determine the interference suffered by the other six mobiles. You can approximate the distance between mobiles with the distance between the cells they are in. Comment on your results. What are the implications of this in termobile terminal of power control?

6.5 The energy per bit in a transmission system is 10 dBm. Using Eq. (6.2), Determine the energy efficiency of error correction codes that have a code rate of 1/5, 1/3, and 1/2 respectively if there are NO errors on the channel. Next consider a harsh radio channel where a stop and wait protocol is employed, i.e., packets are retransmitted if they are not received correctly one by one. About 30% of packets are damaged on this channel on average. The rate 1/5, 1/3 and 1/2 codes can respectively repair 80%, 40%, and 10% of the damaged packets. Determine the energy efficiency of the error correcting codes in this case. Assume that all events are independent.

6.6 A not-so-rich hacker uses an old computer and brute force to break into some wireless systems. It takes him 1 ms on average to test a key to see if it is the right one for an encryption independent of the algorithm employed. How long will it take him to break into an IEEE 802.11 system in the worst case? How long will it take him to break into an IS-136 system on average?

6.7 In Problem 6.6, the hacker realizes that the last six bits of the keys used in a private 802.11 LAN are always zeros. In what time can he break into the system in the worst case?

6.8 In Problem 6.6, the hacker manages to buy a second old computer that can test a key in 1.5 ms. With the two computers, in what time can he break into the system in the worst case? Then he upgrades his computer so that he can test a key in 1 microsecond. In what time can he break into the system in the worst case?