CHAPTER

1

Wireless Network Architecture

INTRODUCTION

Wireless networks come in all shapes and sizes. While many aspects of these diverse networks have common foundations, there are several other aspects that differ widely and need separate treatment. It is thus important to classify and sample example wireless networks to obtain a better understanding of the issues discussed in later chapters. We will start with a discussion of where wireless networks fit in with respect to their wired counterparts. Then we will consider a taxonomy of wireless networks and then consider specific examples of such networks. In the process, we will illustrate the challenges and issues in designing, analyzing, deploying and managing these wireless networks.

1.1 TAXONOMY AND TERMINOLOGY

We start by considering the place of wireless networks in a general networked architecture and then provide one classification of wireless networks that takes into account the topology and coverage of the network.

1.1.1 WHERE DO WIRELESS NETWORKS FIT?

The primary goal of networked systems is to enable connectivity between different devices. The underlying network enables connection of computers to computers, sensors to computers, cell phones

to telephones, computers to telephones and so on. Wired networks have usually been prevalent prior to the corresponding wireless networks. Backbone networks primarily make use of fibre, although microwave point-to-point links and satellite links have been used for long distance communication. So there is already this *fixed network* of devices and links that carries data from some source to the destination, perhaps separated across continents. This network can be the public switched telephone network (PSTN) or the Internet although the boundary between the two has blurred significantly. We will just call this the "fixed network" for now.

When we consider wireless networks today, we are not very interested in the point-to-point wireless links. What makes wireless networks interesting is the fact that they offer *portable* or *mobile* access to networks and applications. By portable, we mean the ability to carry the device *anywhere* and still be able to connect. By mobile, we mean the ability to stay connected at *anytime* even if the user is moving with the device at speeds of up to 100 kmph. Another type of wireless network of interest is one that is made up of *low-power* devices with small form-factor that may be fixed (such as sensor networks or radio-frequency identification (RFID) tags). Note that portability and mobility also mean that devices operate on battery power and have a reasonably small form-factor.

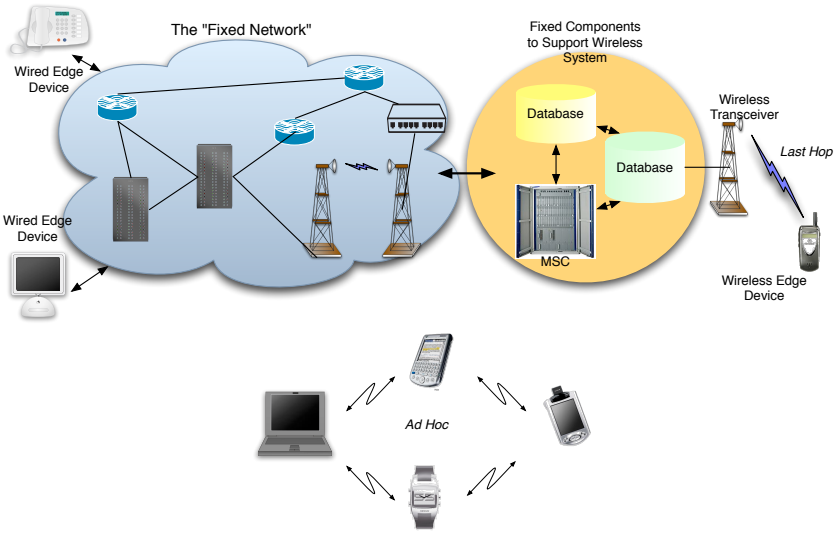

Figure 1.1: Where wireless networks fit in general networked networks

We are usually interested in two types of networks - the last hop and ad hoc clusters. The *Last Hop* wireless networks consist of those that use a wireless link to connect to the fixed network and thereby to other wired or wireless devices that are also connected to the fixed network. In this case, the source or destination or both are wireless devices. They use a *single* wireless link to connect to the fixed network. This also happens to be the first link if the device is the source or the last link if the device is the destination. *Ad Hoc* wireless networks do not typically connect to the fixed network, although there may be some limited connections. In this case, wireless devices connect

to one another directly without an intervening fixed network component. Figure 1.1 shows how wireless networks fit into the big picture of networked networks.

In the top of Figure 1.1 a representative last-hop wireless network making use of some fixed components for support is shown. Sometimes, this is called an *infrastructure topology* as described later. The fixed components shown include some databases and a *mobile switching center* (MSC) that will also be described later. These components are necessary to enable smooth functioning of the wireless network. Note that these fixed components themselves are connected to the fixed network. Cellular telephone networks, most WiFi networks, fixed broadband wireless networks like WiMax are examples of last-hop or infrastructure wireless networks. Not all of them have the entire array of fixed components for support, but have some of these components. We will describe these components and their importance in later sections of this chapter.

In the bottom of Figure 1.1, an ad hoc wireless network is shown where four wireless devices are communicating with each other directly. It is possible to have a WiFi network with purely ad hoc connectivity. Other examples of ad hoc wireless networks are those based on Bluetooth, sensor networks, RFID networks and vehicular communication networks. As there is a lack of fixed components connecting to a fixed network, the wireless devices themselves will have to take up a significant amount of the functionality necessary for supporting smooth operation in the case of ad hoc wireless networks. It is sometimes possible to have a hierarchy of wireless devices to implement this support better.

1.1.2 CLASSIFICATION

We have already qualified some types of wireless networks above by using adjectives such as "portable", "mobile", "last-hop", "infrastructure", and "ad hoc". Now, we will try to formally come up with a classification of wireless networks with specific example technologies and standards.

We first consider the end wireless device. All devices like cell-phones, laptops, sensors, RFID tags, and other devices that use wireless transmissions will be called *mobile stations* or MSs with exceptions where the context requires us to identify the device exactly (for example as a sensor). Note that some of these devices may not be mobile, but we will still call them MSs. Next we consider Figure 1.2 where wireless networks have been classified into three broad categories based on their *topology*.

An *infrastructure* topology is what we primarily encounter today. Here, there is a fixed radio transceiver that connects to the fixed network. MSs in an infrastructure topology have to connect to each other or to other entities in the fixed network using this fixed radio transceiver. It is common to refer to this fixed radio transceiver as a *base station* (BS). There are other names used for the fixed radio transceiver. It is called an *access point* (AP) in WiFi because it is the point of access to the fixed network. A general term used in some standards is "co-ordination point" (CP). In any case, it is possible for MSs to connect to the network through a given BS over some area near the BS. This area is called a *cell*. The BSs are connected over a wired or fixed network so that they can provide coverage over a given geographical area. Thus, a number of "cells" cover the area and enable connectivity of MSs. Wireless networks with infrastructure topology are the most well-developed

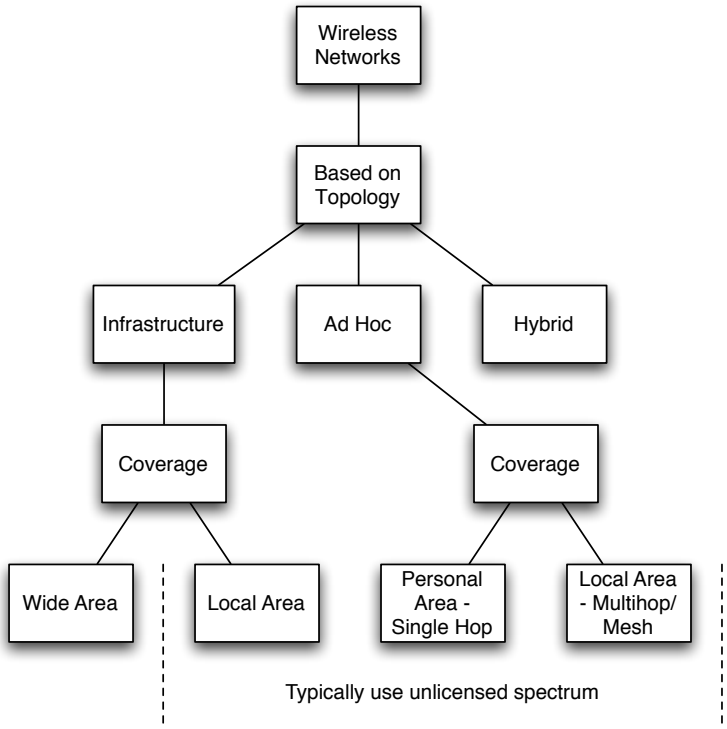

Figure 1.2: A taxonomy of wireless networks

with specific applications,

An *ad hoc* topology is common with personal digital accessories like digital cameras, printers, headsets, cell-phones and the like that connect to one another as and when needed. For example, a cell-phone may connect to a headset when a call is made and disconnect from it after the call is completed. It may connect to a computer at a later time to synchronize calendars and address books. The computer itself may be simultaneously connected to a digital camera that is uploading its pictures. The word ad hoc implies that the wireless network has been formed or arranged for a particular purpose only. Once the purpose is completed, a MS may not be part of the network anymore. Ad hoc topologies are also common in the case of sensor networks. A number of sensors may be randomly deployed to perform some activities over a given area. They may need to communicate with one another as needed.

A *hybrid* or *heterogeneous* topology is a mixture of infrastructure and ad hoc topologies. For example, an ad hoc network of MSs may be created for emergency purposes in a disaster area. Some of these MSs may also be able to simultaneously connect to a cellular network forming a bridge between the ad hoc network and the fixed network. Sometimes, the words hybrid and heterogeneous are used to indicate networks with mixed technologies such as a cellular telephone

network and WiFi network [42]. In such networks, a MS should be able to seamlessly roam from a WiFi network to the cellular network or vice versa.

If we move a level down the classification shown in Figure 1.2, we see that both infrastructure and ad hoc networks can be further classified based on *coverage*. Coverage can be loosely defined as the contiguous geographical area where a MS can communicate with the wireless network of interest. If the geographical area is as large as a city, the wireless network is called a "Wide-Area" network. An example of a Wireless Wide Area Network (WWAN) is the cellular telephone network that provides coverage over cities and sometimes across states. For geographical areas smaller than a city, but large enough to include a building or several buildings in a campus, the wireless network is called a "Local Area" network. Wireless Local Area Networks (WLANs) based on WiFi are examples of local area networks. A "Personal Area" network is one that exists in a small space around a person and his devices. Usually, the communication is only across a *single hop*. That is, a device only communicates with another device if it is in range. For example, a cell-phone connecting to a computer and a headset via Bluetooth creates a Wireless Personal Area Network (WPAN). A device typically does not *relay* packets or messages between two other devices by acting as an intermediary. A second type of ad hoc network allows relaying of packets by devices across *multiple hops* so that the effective communication area is larger (for example a local area). Note that ad hoc networks that cover local areas have been clubbed into the same category as "mesh" networks [15]. The reason for this is that ad hoc networks are neither widespread nor standardized although prototype implementations and working sensor networks do exist. We will revisit this topic later. A last comment on Figure 1.2 is on the use of *licensed* and *unlicensed* spectrum by wireless networks. Most local and personal area networks use unlicensed spectrum with some regulations on the transmit power levels and fair usage. WWANs on the other hand use licensed spectrum that is allocated to the service provider by the Federal Communications Commission (FCC) in the US or another regulatory authority elsewhere.

There are other classifications of wireless networks that are possible - based on the mobility of devices and based on the power consumption of devices. We have already talked about the difference between portability and mobility. Sometimes, fixed and stationary wireless connections are also included in this mix. Power consumption is usually proportionally related to the coverage - a device needs higher transmit power to communicate over larger one-hop distances. Transmissions in WWANs require higher power than those in sensor networks with WLANs being somewhere in between.

1.2 Example Wireless Network Architectures and Operational Issues

In this section we will consider at least one example of each type of wireless network described by the leaves of the tree in Figure 1.2. Our goal in this section is to demonstrate, at a very high level, the operation of some of these networks and introduce the terminology used in these networks. These terms are large in number and will be used in subsequent chapters. The reader is warned about the potential cause for confusion with these terms later on. We will also consider at a high

level, the issues and challenges specific to each type of wireless network since they can appear to be substantially different across the different network types and yet be similar in nature. At this stage, we will not discuss issues such as bandwidth or reliable data transfer as they are essential for all networks.

1.2.1 Cellular Telephone Networks

A cellular telephone network is an example of a WWAN with infrastructure topology. It also by far has the most complex architecture of all wireless networks. Let us consider a general architecture, not specific to any particular cellular telephone network standard, shown in Figure 1.3. We can break up the network into three parts - *the radio subsystem*, *the network subsystem* and *the management subsystem*. While the names by themselves suggest the components of the network, we will look at some details of the subsystems below.

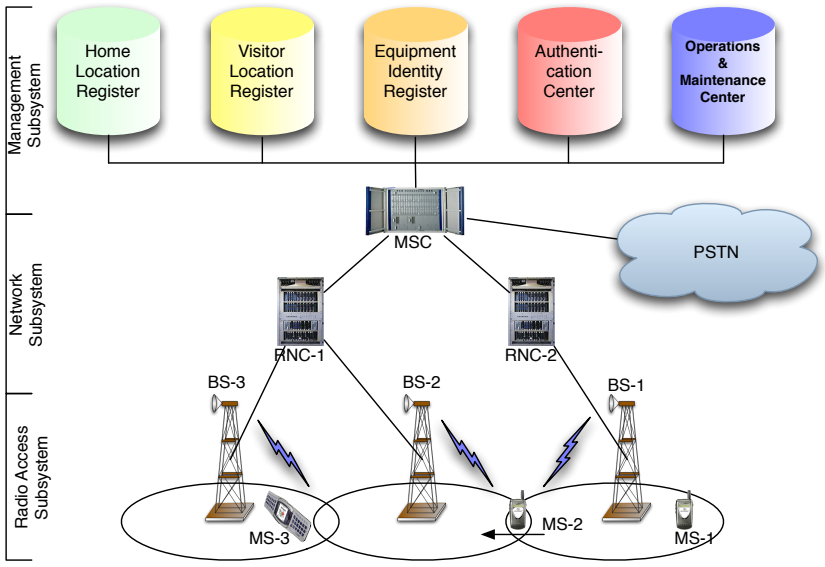

Figure 1.3: A general cellular network architecture

The radio subsystem comprises of the entities in the network that have a *radio interface* or *air interface*. From Figure 1.3, the two components in the radio subsystem are the MS and the BS. The BS provides coverage in a given area called a cell as mentioned earlier. Cells are shown as ovals in Figure 1.3 although they are really irregular in shape as we will see in Chapters 3 and **??**. When a MS powers up, it needs to determine the availability of services in the cell in which it is located. This process is called *service discovery*. For this purpose, all BSs periodically (or continuously) transmit some kind of *beacon* signals which have information related to the network on known frequency channels using known signaling schemes. These beacon signals have different names in different

standards. MSs will first decode this information. If there are multiple beacon signals that are detected, algorithms within the MS will allow it to pick one of them based on one or more criteria such as communications quality, type of network, network capabilities, network ownership, and so on.

In any given cell, as this is an infrastructure topology, MSs always communicate with a BS whether the other party is another MS or a telephone connected to the PSTN. The communication between a MS and a BS is two-way (duplex). The MS can receive and send at the same time, usually by using two different frequencies. The transmission from the MS to the BS is called the *uplink* or *forward channel*. The transmission from the BS to the MS is called the *downlink* or *reverse channel*. Note that multiple MSs connect to the same BS. So a BS must be capable of handling many simultaneous uplink and downlink transmissions. Different MSs may be transmitting and receiving at the same time. These transmissions must not interfere with one another. Separation of these transmissions is achieved by using different frequencies, time slots or spreading codes for each MS. Also, there will be simultaneous transmissions in adjacent and neighboring cells (e.g., cells served by BS-1 and BS-2). To ensure that these transmissions do not interfere, similar steps must be taken. It is common to use different frequencies or spreading codes in different cells, but not different time slots as it is harder to synchronize transmissions. Moreover, the propagation delays are unpredictable since they depend on the locations and physical separation between two MSs.

A technical challenge in the radio subsystem is the dynamic assignment of frequencies, time slots and/or spreading codes to MSs. MSs can begin a call at any point of time and it does not make sense to statically allocate a frequency, time slot or spreading code to a MS. Sometimes, it is possible that certain frequencies are facing more interference resulting in poorer communications quality. In such a case, a MS may be asked to switch to a different frequency or increase its power. The management of the time slots, frequencies, spreading codes and transmit power of a MS (and the corresponding changes at the BS) is typically the task of the *Radio Network Controller* (RNC) which is part of the network subsystem. The time slots, frequencies, spreading codes, and transmit power are the so-called *radio resources* that need to be efficiently managed. Hence this task is called *Radio Resources Management* (RRM). An RNC controls many BSs and performs allocation of the radio resources through the BSs. In Figure 1.3, notice that MS-2 is moving away from BS-1 towards BS-2. As it moves in this direction, the communications quality of the link between MS-2 and BS-1 deteriorates and at some point of time, MS-2 must switch to BS-2 as the point of access to the network. This is called *handoff*. The decision of performing the handoff requires using metrics associated with the radio resources and is part of RRM. Note that measuring the metrics associated with radio resources (communications quality, received signal strength, and so on) is not a trivial task either. Often, the beacon signals are used as reference signals for comparison and determination of radio resource metrics.

For communicating with the external world, a *Mobile Switching Center* (MSC) is required. The external world includes the PSTN and other cellular telephone networks. One MSC usually controls a group of RNCs creating a tree-like hierarchy. Usually, the BS by itself is only a transceiver. To set up a communication between a MS and another party, the RNC and the MSC need to intervene. For example, if the MS is communicating with a telephone connected to the landline PSTN, the

MSC needs to communicate with switches in the PSTN and set up a circuit between the last switch in the PSTN and the MS. This circuit will go through the RNC controlling the BS over the air to the MS. You can immediately see two problems. First, let us suppose the MS is not connected (it is idle or on stand-by) and someone somewhere wants to place a call to this MS. How does the network know where the MS is located even if it is powered up? To solve this problem, *location management* techniques are used in the network. A MS can use the beacon signals to determine whether it is in the same location area (essentially a group of cells) as it was sometime back or whether it has moved. In any case, it sends a location update message periodically to the network. These location update messages are carried to a database called the *Visitor Location Register* (VLR) that is connected to the MSC. Note that the VLR and the other databases discussed below are part of the management subsystem. When the MS powers up and sends a location update message, the VLR contacts another database called the *Home Location Register* (HLR) and exchanges this information with it. The HLR keeps a pointer to the VLR that is serving the MS currently. When a call arrives, it is first sent to a MSC connected to the HLR. The HLR uses the pointer to redirect the connection to the MSC connected to the VLR serving the MS. This MSC uses the RNC to *page* the MS in a group of cells where it was last reported to be and obtain a response.

Second, if the MS is already connected to another party and moves from one BS to another, how is the circuit maintained? To solve this problem, *handoff management* techniques are used in the network. The VLR and HLR exchange information about a MS as it moves keeping track of it. If the MS crosses to a new BS that is controlled by the same RNC as the previous BS, no changes need be made except for the RNC to set up a circuit between itself and the new BS. If the new BS is controlled by a different RNC, the MSC will have to get involved. If the handoff takes place to a BS that is controlled by a different MSC, a new VLR will have to get involved in migrating the circuit.

MSs not in active use will often go to sleep or stand-by to save battery life. They wake up occasionally to see if there are any pages from incoming calls for them. The network and MSs have to maintain some synchronization and clocking to ensure that the MS indeed sees any messages intended for it when it wakes up. Protocols have to make use of this timing information to ensure reliable delivery of information as well.

When the MS first powers up and tries to connect to the network based on some set of decoded beacon signals, an *authentication* and *key establishment* process needs to take place. This will ensure that the MS is authentic and will create keys in the MS and the RNC that will allow communications to be encrypted. The cryptographic keys used for this purpose are known only to the MS and a database called the *Authentication Center* (AuC). The AuC communicates temporary secret information to the MSC/RNC/BS to enable the authentication of the MS and encrypted communication between the entities. The *Equipment Identity Register* (EIR) is a database used to verify whether the MS is legitimate (not stolen, cloned or one that has not paid the subscriber fees). The *Operation and Maintenance Center* (OMC) handles billing, accounting (e.g., roaming, peak minutes and so on) and other operational tasks.

Clearly, you can see that the cellular telephone network is a fairly complex network of many different entities handling many different tasks. In fact, this is by far the most complex wireless network. As we will see next, the IEEE WLAN is not as complex in its architecture although it has to perform almost the same number of tasks as a cellular network.

1.2.2 Wireless Local Area Networks

In this section, we will see how WLANs operate in general. Once again, we will not try to consider standard-specific operation, but try to provide a general overview. First, we will consider an infrastructure topology shown in Figure 1.4. Notice that this architecture looks much simpler than the corresponding cellular network architecture in Figure 1.3.

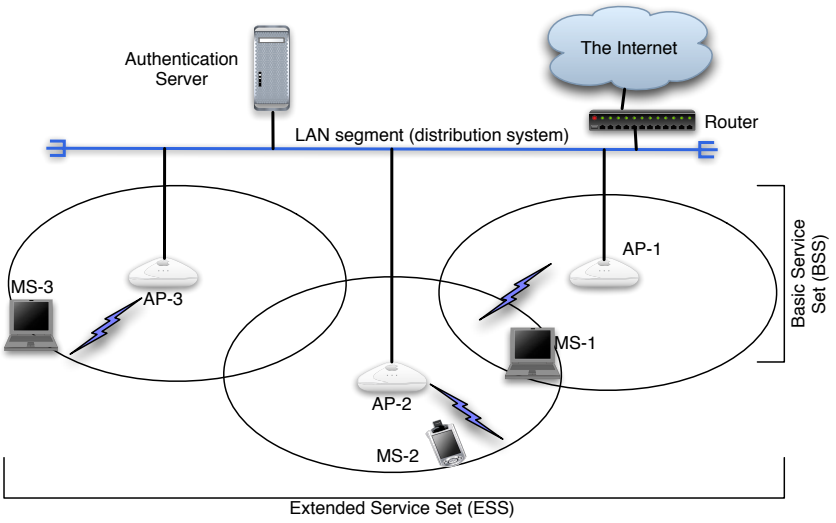

Figure 1.4: A general WLAN architecture

Like the cellular network, we can think of a radio subsystem in WLANs comprising of MSs and BSs (that are called access points - APs). The APs connect to the fixed Local Area Network (LAN) which connects to the Internet through a router. APs cover areas called cells (which are also irregular in shape although shown as ovals in Figure 1.4) and transmit beacons periodically to enable service discovery. Such periodically broadcast packets are in fact called *beacons* in the WiFi standard and contain information about the *basic service set* (BSS) and the *extended service set* (ESS). The BSS is one AP, the cell it covers and all MSs connected to it. The ESS is the set of all APs on the same network and all the MSs connected to them. Note that APs in the same network are connected through a wired LAN segment as shown in Figure 1.4. The wired network that connects the APs is called a *distribution system*.

In a LAN, packets are *broadcast* on the medium and all devices connecting to a LAN share the medium using a *Medium Access Control* (MAC) protocol. Devices on the LAN pick up packets if they are addressed to them and discard them otherwise. Note that entities such as the RNC and MSC do not exist in the WLAN architecture. Moreover, the AP and MSs have to be inexpensive. Thus, the MAC protocol has to be simple and distributed so as to not require central control or expensive components. Thus, there is no concept of having different frequencies or time slots to separate

transmissions. MSs and APs transmit on the same frequency channel, but these transmissions have to be somehow kept separate. Carrier sensing is a common methodology used in LANs for medium access. Loosely speaking, each MS senses the medium to see if there is a transmission. If the medium is free, the MS can transmit its packet. The actual process is more complex because the MAC protocol has to handle *collisions* that occur when two MSs transmit at the same time.

RRM is minimal in a WLAN because there is really no need for allocation of time slots, spreading codes or frequency carriers. There is no transmit power control - retransmissions take care of collisions or packets lost due to poor communications quality. Determination of the need to handoff does exist. The MS makes a decision to handoff if it sees poor signal quality which is measured on beacon signals. Location management is simple. Every device on a LAN is supposed to receive packets and it is up to the device to take a packet or discard it. So all a sending party needs to know is the IP address of the destination device on a LAN. An IP packet destined to the device reaches the LAN through a router attached to the LAN. The packet is simply put on the LAN segment and the device will pick it up. In the case of a MS, things are a bit different. A MS registers with an AP on the LAN. The AP has to pick up the packet intended for the MS and transmit it on the air. All MSs in the BSS receive the packet. Only the intended MS will pick up the packet and the rest will discard the packet.

APs communicate through the distribution system when a MS hands off from one AP to another. There is really no complex handoff management scheme. What happens if a MS moves *across* LANs so that its IP address changes? There are two possibilities. The client applications on the MS have to handle the change in IP address. This is easy for applications like web browsing, e-mail, and new instant messaging sessions. Mobile IP is required for ongoing communication sessions (e.g., a voice over IP call or on ongoing instant messaging chat session).

When a MS powers up, it performs an association with an AP. It picks the AP based on the quality and content of the beacon signal. Association with an AP can be secure or not. A secure association requires authentication followed by encryption. Keys in older protocols were manually installed in the AP and MSs and the authentication was performed at the AP. More recently, it is possible to have an Authentication Server (AS) that can be used for authentication and key establishment similar to cellular networks (except that the AS is usually located in the same network). All communications on the air are subsequently encrypted. Management and operations are performed extraneously in a manner similar to the management of wired computer networks.

1.2.3 Wireless Personal Area Networks

WPANs originated through their history in two different applications. Bluetooth, an example of a WPAN technology were initially envisaged as a "cable replacement" technology. Another application of interest was the so-called "BodyLAN" or "Wearable Ad Hoc Network". Together, they were collapsed into the idea of a WPAN [18] which is an independent wireless network of devices in a *Personal Operating Space* (POS) with radius 10m around one or two human beings. The MSs in this network are all within the communication range of one another (see Figure 1.5). A network is created spontaneously and unobtrusively as and when the MSs need to communicate with one another.

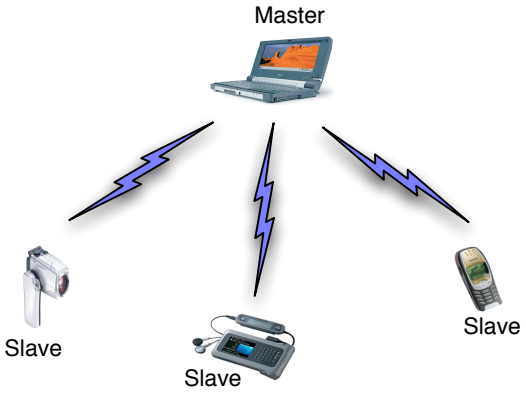

Figure 1.5: A general WPAN architecture

As the network is supposed to be plug-and-play, self-configuration and service discovery become important issues. Recall that there is no central fixed transceiver (such as a BS or AP) in this network. So MSs have to discover one another and also discover the capabilities of one another. This is done using specific self-configuration mechanisms and service discovery protocols that we will discuss in later chapters. In the case of Bluetooth, a Master device initiates communications and polls all responding Slave devices for data as and when required (see Figure 1.5). Most often, MSs operate using a single available frequency band for communications. Interference and coexistence with other networks (WPANs or WLANs) is a significant issue here. To avoid interference between co-located WPANs, frequency hopping within this band is adopted by Bluetooth. In the 802.15.4 WPAN standard, carrier sensing is employed as in the case of WLANs. A very important requirement in WPANs is to have the ability to conserve power to the maximum possible extent. MSs in WPANs should be able to operate for days. Mechanisms are also necessary to authenticate MSs to one another and establish keys for encrypted communications between MSs. In Bluetooth, this is achieved by installing PIN numbers in trusted MSs. The PIN numbers along with the MAC addresses and other random numbers are employed for key establishment.

Location management, handoffs, and RRM are not as important in WPANs as they are for cellular networks, especially because of the medium access control protocol and low transmission power.

1.2.4 Mobile Ad Hoc Networks

The definition of Mobile Ad Hoc Networks (MANETs) slightly modified from a description in the online encyclopedia Wikipedia [12] is as follows:

> "A mobile ad-hoc network is a self-configuring network of mobile routers (and associated hosts) connected by wireless links – the union of which form an arbitrary topology.