

Unit 1:

Wireless Network Technology

The changes in wireless technology for data networks over the past five years have been more dramatic than the changes in radio itself in the century since Guglielmo Marconi sent the first wireless telegraph signal across the Atlantic from Cornwall, England, to St. Johns, Newfoundland, on December 12, 1901. The progress in commercial radio transmission from telegraphy to voice to television was measured in decades. Commercial digital wireless transmission began in the mid-1990s when cellular digital telephony—known as PCS for Personal Communications Service—replaced Advanced Mobile Phone Service (AMPS), the then-dominant analog voice transmission protocol. Digital wireless telephony technology was then split into two competing technologies: time division multiple access (TDMA) and code division multiple access (CDMA). Older TDMA has now been replaced by Global System for Mobile Communications (GSM), a standard version of TDMA used by most European and Asian carriers as well as by some North American cell phone carriers. CDMA is used by some Japanese carriers as well as by two North American cell phone carriers. TDMA, GSM, and CDMA are not interoperable due to major protocol differences and frequency assignments.

As the cell phone industry has evolved, a much greater capability for digital data transmission has become necessary, particularly to support “smart phones.” Several different protocols have been tried, but the industry is now converging on LTE (Long Term Evolution), a protocol closely related to GSM. Even carriers using CDMA for voice are using LTE to achieve “4G” network speeds for digital data.

The wireless local area network (LAN) began to emerge in the late 1990s, when it became obvious that there was a need for wireless data networking. Wireless LANs required faster data transmission than was possible with cellular PCS (of any technology), and eventually suppliers settled upon using digital spread-spectrum as defined by the IEEE 802.11 standards. Spread-spectrum was originally developed for the U.S. military so wireless transmissions could be made in the

presence of strong jamming signals. This work by the military was based on the spread-spectrum patent U.S. 2,292,387, which was originally granted to Hollywood actress Hedy Lamarr and her partner George Antheil.

Frequency hopping spread-spectrum (FHSS) and direct sequence spread-spectrum (DSSS), both operating up to 2.0 Mbps, were the first two IEEE 802.11 technologies. Neither is commercially available today. These initial technologies were improved upon until, in rapid succession, IEEE 802.11b (operating at up to 11 Mbps) and 802.11a, as well as 802.11g (both operating up to 54 Mbps) emerged. All of these are called Wi-Fi (wireless fidelity) after the name of the supporting industry association, the Wi-Fi Alliance, but 802.11b (and later 802.11g) became commercially successful technologies with a large home and office installed base. Both 802.11n and 802.11n/dual frequency have now essentially displaced 802.11g and IEEE802.11a. For the sake of simplicity, I will continue using the term Wi-Fi to represent any of the IEEE 802.11 standards.

1.1 Wireless Standards

The dynamic nature of wireless digital data communications stems from the standards committees of the Institute of Electrical and Electronic Engineers (IEEE), which develops most of these protocols. No illusion currently exists within the IEEE 802 committee, which is responsible for personal area networks (PAN), local area networks (LAN), and metropolitan area networks (MAN) that it will be possible to create a single network protocol useful over all of these four domains. Therefore, each application that has a special interest that is not accommodated by an existing protocol can form a new subcommittee to create a new protocol standard. IEEE ensures only that these subcommittees deliberate fairly, that they do not exclude a genuine interest, and that all proposed standards are publicly reviewed. All IEEE 802 standards are automatically submitted to the ISO/IEC (International Organization for Standardization and International Electrotechnical Commission) for consideration as international standards. Several of the IEEE 802 standards have failed in the marketplace, while others have succeeded.

In their efforts to make the family of IEEE 802 standards general, and independent of applications, all 802 committees develop their standards to define only Layers 1 and 2 of the ISO seven layer protocol stack (IEC/ISO 7498-1) shown in Table 1. By standardizing at this low level, many application-oriented upper layers can be added by standards committees and consortia focusing on specific

application areas to take advantage of the commodity volume market for chips based on the IEEE 802 protocols.

Table 1. IEC/ISO 7498-1 Seven Layer Stack

Layer No.	Name	Function
7	Application	User interface
6	Presentation	Data format conversion
5	Session	Connection
4	Transport	Acknowledged service
3	Network	Network Addressing
2	Data Link	Local address and mastership
1	Physical	Hardware/Spectrum

Another source of wireless communications protocols is the International Telecommunications Union (ITU), the standards body for telephone networks. Telephony has rapidly changed from a purely wired circuit-switched analog service—also known as POTS (Plain Old Telephone Service)—to VoIP (Voice over Internet Protocol) on broadband service and now to wireless telephony.

Wireless telephone service began with AMPS, and rapidly progressed to 1G (first generation) digital service, which evolved to 2G, 2.5G, then to 3G and now 4G. Digital wireless telephony implementation has not been uniform, with a split between TDMA (especially since the GSM version of TDMA has been adopted in most countries) and CDMA, available in fewer countries.

Even though there are excellent reasons to keep CDMA, GSM, 3G, and 4G wireless in mind for in-plant voice networks and some mobile data applications, they are not presently being considered for industrial use. However, given the eventual availability of low-cost and low-power consumption 3G and 4G technologies, they should not be ignored.

A word of caution is in order about the standards documents for data communications. These very large documents are not intended to be read by general users.

They are written for the implementer of networks and networking devices. If you really want to see some examples of such standards, however, most IEEE 802 network documents more than six months old are available for download on the IEEE standards website: <http://standards.ieee.org/getieee802/> (look for the *click here* link in the last paragraph of the text). The IEEE and others often publish books about the standards, making them easier to understand.

1.1.1 IEEE 802.11, Wi-Fi

One factor causing rapid technological change in wireless communications is the ever-increasing capacity of commercial semiconductor processes such as CMOS (complementary metal oxide semiconductor) to handle higher frequencies. This factor alone is responsible for the recent rise in interest in IEEE 802.11n/dual frequency, which previously required more expensive GaAs (gallium arsenide) processes or higher-power bipolar semiconductors to use the 5-GHz frequency band. When 802.11n or dual frequency 802.11n chips are built in CMOS, they are as economical as the slower 802.11g chips.

With the ratification of 802.11n and the subsequent flood of new products on the market, we have witnessed another dramatic change in the Wi-Fi market. 802.11n, which is backwards-compatible with 802.11g, has completely displaced 802.11g for new wireless products. As the Wi-Fi band at 2.4 GHz becomes saturated, the benefits of dual frequency 802.11n become compelling since the 5.0-GHz band offers 24 non-overlapping channels, versus just three for service at 2.4 GHz. Chips that offer dual frequency 802.11n are already on the market, and soon all new Wi-Fi LANs will offer access at both 2.4- and 5-GHz service at little to no price premium. Additionally, IEEE 802.11ac chips that operate only in the 5-GHz band are also commercially available. Figure 1 illustrates a roadmap for these transitions in the Wi-Fi market.

The most appealing technology embedded into 802.11n, in addition to its dual frequency band, and in IEEE 802.11ac is called MIMO (Multiple Inputs, Multiple Outputs), most easily recognized by several antennas on dual frequency 802.11n and 802.11ac products. The “n” and “ac” standards require that all signals be simultaneously transmitted on each of the antennas. Due to the spatial separation of these antennas (they are a few centimeters apart) signals transmitted by all antennas will be received by the multiple antennas of the receiver slightly out of phase with each other. MIMO technology provides a way for the receiver to mathematically realign the phases of the received signals such that

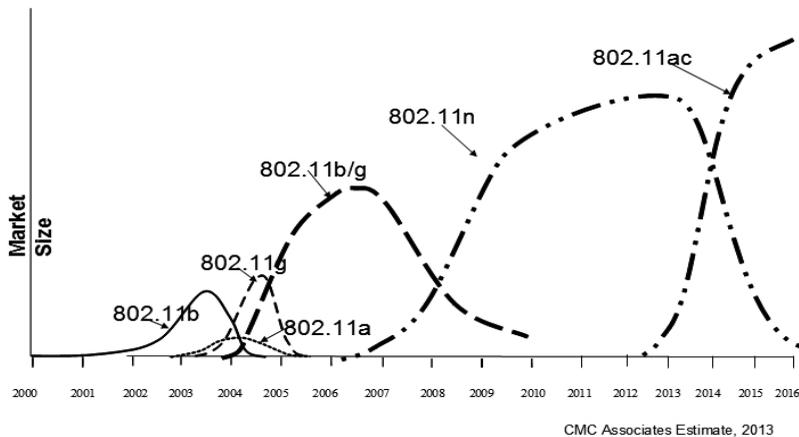


Figure 1. Market Size for Versions of Wi-Fi

the resulting resolved signal is now stronger and more reliable than any single signal. Note also that reflected signals, often called multipath signals, are also out of phase with the original. MIMO offers a technical solution to the multipath problem, which is often associated with networks built in large plant units in the process and metals industries, often referred to as the “canyons of steel.”

IEEE 802.11n can also bond channels in both the 2.4- and 5-GHz ISM (Industrial, Scientific, and Medical) bands that were formerly assigned to 802.11g and 802.11a, respectively. Channel bonding in 802.11n may be used to achieve a higher data rate. While a single channel for either “a” or “g” can achieve a theoretical maximum 54 Mbps, bonding two channels in 802.11n can achieve a theoretical rate of 108 Mbps. 802.11n can achieve theoretical data rates as high as 600 Mbps by bonding four channels.

Channels may only be bonded within bands. This means that currently, an IEEE 802.11n dual-channel device requires two radios, one for each band. Many inexpensive 802.11n devices may not be able to implement the dual radio part of IEEE 802.11n simply because they do not have a 5-GHz radio, and as a result, will not be able to achieve the higher data rates that can come from channel bonding.

The successor to 802.11n in the Wi-Fi market is IEEE 802.11ac, which is very much like IEEE 802.11n except for its ability to bond more channels in the 5-GHz band to achieve a maximum data rate greater than 1.0 Gbps. This becomes

a wireless equivalent of Gigabit Ethernet. While the standard has not yet been completed (as of 2013) there are already products on the market. By the end of 2014, IEEE 802.11ac should be the dominant Wi-Fi protocol.

Table 2. IEEE 802.11 Standards

Standard	Release date	Band (GHz)	Bandwidth (MHz)	Theoretical Max Data Rate	Advanced Antenna Technologies
802.11	1997	2.4	20	2 Mbps	N/A
802.11b	1999	2.4	20	11 Mbps	N/A
802.11a	1999	5	20	54 Mbps	N/A
802.11g	2003	2.4	20	54 Mbps	N/A
802.11n	2009	2.4, 5	20, 40	600 Mbps	MIMO, up to 4 spatial streams
802.11ad	2012	60	2160	6.76 Gbps	Beam forming
802.11ac	2013	5	40, 80, 160	4.9 Gbps	MIMO, MU-MIMO, up to 8 spatial streams

Courtesy of IEEE 802 standards committee

Another recent standard is IEEE 802.11ad (2012), which deviates from the 2.4-/5-GHz ISM band by using an ISM band at 60 GHz that currently is lightly used. It uses multiple antennas to form a virtual “beam” between the transmitter and receiver. This makes IEEE 802.11ad almost as effective as a directional antenna such as a Yagi (Figure 2) but not quite as effective as a parabolic antenna (Figure 3). The beam forming is performed with advanced mathematics setting both the amplitude and phase shifting of the signal sent to and received from each of the up to eight antennas. This is an implementation of the phased array antennas discussed in section 1.4.5. By using the 60-GHz band, sufficient bandwidth can be allocated to allow data rates up to 6.67 Gbps, which is known as WiGig. It is not intended for conventional LANs, but is directed at cable replacement for high definition video and audio signals.

As mentioned above, the Wi-Fi market is supported by the Wi-Fi Alliance, which in its own words is “a nonprofit international association formed in 1999 to certify interoperability of wireless Local Area Network products based on IEEE 802.11 specifications.” The Wi-Fi Alliance currently has more than 200

member companies from around the world, and more than 1000 products have received Wi-Fi[®] certification since certification began in March of 2000. The goal of the Wi-Fi Alliance's members is "to enhance the user experience through product interoperability." The Wi-Fi Alliance website is: <http://www.wi-fi.org/>

1.1.2 IEEE 802.15.1, Bluetooth

Bluetooth is a wireless communications technology that was created to eliminate the wires connecting cellular telephones and other portable electronics to a headset or earphones. Bluetooth has already been applied in many commercial products and almost all wireless telephones, but at a much slower pace than its developers ever dreamed. Bluetooth has just enough networking capability to interest a wide variety of companies in extending its use beyond its original scope.

In fact, Bluetooth is far more than a communications protocol; it is also a full communications application stack. The lower two communications layers of Bluetooth (Layers 1 and 2) have been published as IEEE standard 802.15.1. For the original task of device connection, the upper layers of Bluetooth offer a rich suite of functionalities, including enabling walk-up linking without user interaction and establishing voice connection.

Bluetooth networking is intentionally limited to a maximum of eight Bluetooth nodes, which together form a *piconet*. When a node is included in more than one piconet, that node then assumes the routing task of forwarding messages to/from the other piconet, adding a form of mesh networking to the complexity of Bluetooth. The most attractive feature of Bluetooth for industrial automation purposes is its use of forward error correction (FEC) for delivering messages without error and without requiring retransmission. The drawback of FEC is loss of efficiency: a 1-Mbps communications channel can deliver only 721 Kbps.

A multivendor consortium (Bluetooth Special Interest Group [SIG]) defined Bluetooth, not a standards organization. With the consent of the SIG, the lower two layers of Bluetooth were reformatted and have now become the IEEE 802.15.1 standard. Just like 802.11b and 802.11g, Bluetooth operates in the unlicensed 2.4-GHz frequency band, but it uses frequency-hopping spread-spectrum technology that hops faster than the original FHSS of 802.11. As a result, the presence of Bluetooth nodes in close proximity to Wi-Fi nodes causes the signal for the WLAN to degrade, sometimes spelling disaster for Wi-Fi transmissions.

Bluetooth 1.2 and later protocols help such nodes avoid Wi-Fi signal degradation by listening for signals on the radio channels before transmitting. Many early suppliers of nodes with both Bluetooth and Wi-Fi capability have been able to synchronize transmissions to avoid degradation. Suppliers of 802.11a and 802.11ac devices, which operate in the 5-GHz unlicensed band, are quick to point out that they avoid signal degradation from Bluetooth completely. Nevertheless, 802.11g and the single frequency of 802.11n suffer the same potential problems in the presence of Bluetooth.

Bluetooth Low Energy (BLE)

Part of the Bluetooth 4.0 specification is a low energy form of Bluetooth communications designed to operate within the smartphone environment and other applications requiring longer battery life than can be sustained by the original Bluetooth specification. BLE operates in the same 2.4-GHz band, but uses direct sequence spread-spectrum rather than frequency hopping. Power is conserved by operating at a reduced duty cycle, sleeping for a majority of the time and awakening only to process data transfers. The data rate is also about one-fourth of the full Bluetooth rate over distances about one-half of the full Bluetooth range. However, there is no significant coexistence problem with Wi-Fi communications.

If you want to know more about Bluetooth, a rich source of information can be found on the official Bluetooth SIG website: <http://www.bluetooth.com/>.

If you want to develop Bluetooth products, the Bluetooth developers' website offers lots of reference material and discussion groups: <https://www.bluetooth.org/>.

1.1.3 IEEE 802.15.4 and ZigBee

Wireless standards have been developed for “personal area networks” or PANs. The most successful of these standards is IEEE 802.15.4, which has served as the base (Levels 1 and 2 of the ISO protocol stack) for ZigBee, ISA100 Wireless, WirelessHART, WIA-PA, and several others. Each of these standards has added its own upper-layer protocols to suit applications in many different markets. Common among all of these networks is a need to consume as little electric power as possible to allow extended battery operation, while covering an area approximately the size of a football field.

Reduced power consumption is achieved by the device's sleeping up to 97 percent of the time, awakening only to send and receive data during a device's "slot time." The protocol defines the timing references necessary to coordinate both network nodes to be awake at the same time: the slot time.

ZigBee builds upon IEEE 802.15.4, adding its own protocol features to meet the needs of several markets: lighting, heating, ventilating and air conditioning, intrusion detection, and factory automation. One of the protocols added by ZigBee is the formation of mesh networks using a Layer 3 protocol. The ZigBee Alliance website is located at www.zigbee.org.

1.2 Proprietary or Non-Standard Wireless Networks

Standards are developed at a much slower pace than the pace of technology. Commercial suppliers often will not wait for the approval of a standard, or may have a product concept that adequately fulfills the network requirements although it does not meet any existing or proposed standard. These companies often introduce their network products in hopes of establishing a market in the absence of a standard. The experience gained by these suppliers can often be useful to the designers of network standards. Sometimes, such a network can become the basis of a standard.

Adaptive Instruments (now a division of Schneider Electric) still offers their own wireless networks for process control field instrumentation. The Accutech network uses frequency-hopping spread-spectrum operating in the 915-MHz ISM (Industrial, Scientific, and Medical) band. This network is capable of passing data at rates that vary from 4.8 to 76.8 Kbps over distances that vary from 780m to 175m. Their devices are battery powered and have a battery life estimated to be several years. An Accutech network is configured with a wired base-station located close to the field instruments, and forms direct links to each instrument from the base station.

Since 2007, Honeywell has been reselling the Accutech instruments, rebranded as their XRY 5000 product line and still supported, but XRY 5000 has been superseded by their ISA100 Wireless product, called OneWireless.

1.3 Process Automation Wireless Networks

There are three standard wireless network architectures that were developed for the industrial process control and data acquisition market: ISA100 Wireless, WirelessHART, and WIA-PA. All three architectures are based on the same radio technology, IEEE 802.15.4-2006. All three modify that radio technology by using channel hopping among the 16 channels specified for IEEE 802.15.4 in the 2.4-GHz ISM band. However, they are not interoperable since they implement some of the basic network functions differently. The following summarizes how these similar networks differ:

- **Sense of time.** ISA100 Wireless uses a time distribution method from the IEEE 1588 standard and that is accurate to ± 62 microseconds. WirelessHART counts slot times, but does not control accuracy. WIA-PA depends on the Network Manager to set time at the remote instruments.
- **Slot time.** ISA100 Wireless has a default of 10 ms slot time, but its slot time can be varied to meet needs. WirelessHART specifies only a 10 ms slot time. WIA-PA depends on the Network Manager to set the slot time.
- **Hopping tables.** ISA100 Wireless has a pre-defined hopping table, and includes the same hopping sequence specified by WirelessHART, which has only its own hopping table. The Network Manager sets the hopping table for WIA-PA.
- **Meshing.** ISA100 Wireless transmits to each neighbor in the mesh during the same slot time using duo-cast protocol. WirelessHART transmits to its neighbors in consecutive time slots. WIA-PA requires that the Network Manager specify and download the routing tables.

1.3.1 Wireless versus Wired Networks

Wi-Fi has generally been considered to be Wireless Ethernet, but it is far more than that. Wired networks, such as Ethernet, are designed for communications between fixed locations. Wireless networks, such as Wi-Fi, are designed for communications between devices. The distinction is lost for fixed-location devices, but device mobility is a primary benefit of wireless. Mobile applications are often found in discrete parts manufacturing and assembly, and in all types of

warehouse applications. However, the primary applications for wireless in industrial automation are expected to be between fixed locations.

The air is free, but to operate, wireless networks often need a wired connection to a computer, a portable terminal, or the wired network along with a source of power and radios. Estimating the cost of a wired network is easy. It is the sum of the cost of the network cable, junctions, and connecting wires; the cable and junction installation; the network interfaces; and the long-term maintenance of the installed wiring plant.

Wireless networks are more difficult to estimate. They include the cost of wiring to access points, access point equipment, wireless interfaces, and long-term wireless troubleshooting and maintenance. While there are fewer items to install and maintain, experience with the installation and maintenance of wireless equipment is currently much more limited than it is with wired.

The other notable problem of wireless devices is that they still need a power source. Wired network nodes can draw power from the local AC receptacle or another local power source, but mobile wireless devices depend on batteries or some alternative power source. Of course, you can often plug the wireless device into a local power source, but then you lose the mobility advantage and incur the cost of installing power connections at the device. To some extent, the PoE (Power over Ethernet) standard, IEEE 802.3af, was created to help resolve part of this problem by supplying electrical power on the wired Ethernet network so it can be used by wireless access points. This standard seems to be well accepted for business or commercial access points. However, PoE still does not address the issue of powering the wireless end-device itself.

1.3.2 Signal Loss/Fading

Now, in the early twenty-first century, wireless networks still suffer from mysterious dead spots—areas where there is no reception. We say “mysterious” because even careful planning cannot remove all dead spots, and sometimes live spots just move or, in the language of radio, fade. The spontaneous loss of communications for no apparent reason is probably one of the most irritating aspects of wireless. The signal often mysteriously returns even before the cause of its loss can be investigated. This occurs with cellular telephones, with Wi-Fi devices, and with all other wireless LAN technologies.

Signal loss can be caused by interference from other radio signals present in the same part of the spectrum as well as by moving equipment. Sometimes, a live spot exists only as a result of a multipath effect when the signal is reflected from some stationary object; sometimes the multipath signal interferes with and cancels the primary signal, causing a dead spot. Wi-Fi also seems to fade in areas in which microwave kitchen appliances are in use or in which a cordless telephone is operating at 2.4 GHz. (Note that the term “cordless telephone” is used to distinguish an ordinary voice telephone from a cellular telephone that is termed a “wireless telephone.”) Actually, in these instances the signal loss is due to interference, which is difficult to distinguish from fading.

Dead spots may occur within buildings, depending on their materials of construction. In the path between the access point and the wireless device, each time the radio waves pass through a solid object the signal is attenuated. Denser materials attenuate the signal more than do less dense materials. Metals used in building construction, particularly steel, may absorb or reflect most of a radio signal, creating a dead spot in the object’s radio “shadow.” Moving the access point or the device by a small amount, perhaps only a few millimeters, may eliminate the dead spot.

Finally, there are sunspots! The sun emits a broad spectrum of electromagnetic waves at all frequencies, which generally constitutes noise. Once in awhile, the surface of the sun experiences flares or dark spots that emit very strong electromagnetic waves that are known to interfere with radio transmissions, and occasionally with wired communications as well.

1.3.3 Multipath Distortion

Radio waves move from an omnidirectional antenna in all directions. When these radio waves strike a very dense material such as metal or stone, they are reflected, much as light is reflected from a mirror or another shiny surface. Even when there is a clear path between the transmitting and receiving antennas, some of the signal that has been reflected along other paths will arrive at the receiving antenna. This phenomenon is called multipath distortion and can distort the received signal since the longer paths will cause the reflected signals to be received out of phase with the signal from the direct path.

The effect of multipath reception can range from nothing to the cancellation of the signal, depending on the paths and the resulting delays.

In some cases, however, the multipath effect can be electronically phase-shifted to boost the received signal. This occurs when multiple paths are received in phase, such as when multiple transmitting antennas are used. In fact, this phenomenon is used by IEEE 802.11n and IEEE 802.11ac. The technology for using the multipath signal to enhance performance is called MIMO (Multiple Input, Multiple Output). MIMO uses multiple antennas on both the transmitter and the receiver to achieve multiple transmissions, and to receive all signals. MIMO circuitry intentionally phase shifts the signals to improve reception.

1.3.4 Shared Airwaves

One of the problems of radio is that the radio frequency spectrum is limited, and new uses are constantly being found for it. The attempt to allocate certain frequency bands for specific uses is the responsibility of governmental agencies—the FCC (Federal Communications Commission) in the United States. The frequency assignment process is highly political and is based loosely on technology. Furthermore, frequency assignment is highly dynamic and is sensitive to economic conditions and the appearance of new technology solutions. For example, the FCC originally assigned 82 6-MHz frequency channels exclusively for broadcast television—an enormous segment of the spectrum for a single purpose. In most areas of the United States, only a tiny fraction of that spectrum is actually being used in any one location. Furthermore, the assigned channels tend to be located in the lower channel numbers, since the higher UHF frequencies have limited distance reception range. Television channels are reused based on geography—when stations are far enough apart to not interfere with each other. Some of the unused UHF television channels have already been reclaimed for other uses, and more are scheduled to be reclaimed in the future since all domestic television in the United States has been converted to digital broadcast.

While only 1 MHz is needed to digitally broadcast the same content that previously required 6 MHz for analog, each broadcaster is still assigned their original 6-MHz band. Much of that 6-MHz band is needed for the broadcast of high definition television (HDTV) simultaneously with the digital version of standard television. Needless to say, television stations are highly reluctant to change frequency channels once they are in use, or to relinquish any of their assigned bandwidth.

One of the controversial uses of the spectrum is called “White Space” (TVWS; television white space). This is the bandwidth that was formerly called the

“guard band” between allocated television channels. This unused portion of the spectrum was formerly unusable with analog television broadcast, but with narrowband digital television, the white space between those 6 MHz channels is now being made available, but the process is not yet completed. Allocation of some of this white space to industrial wireless is included in the discussions.

The United States military is one of the most demanding users of radio frequencies and is very reluctant to give up any frequency previously assigned to it. This same attitude is reflected in the military establishments in most other countries as well, even when the service using that frequency has been abandoned. Another demanding public sector is amateur radio, which has been allocated small frequency bands scattered throughout the spectrum. Amateur radio broadcasters are also reluctant to abandon any frequency band.

Nevertheless, the United States and most other governments have ordered that all allocated users share the radio spectrum unless the service cannot function when shared. By definition, the military frequency bands cannot be shared. Public radio, television, and global positioning satellite (GPS) frequencies also cannot be shared. Certain public safety and many business uses are licensed and are not shared. The remainder can be shared, and they are divided into both licensed and unlicensed frequency bands. Generally, licensed bands allow users to broadcast at higher power ratings in order to reach longer distances, while unlicensed bands are forced to limit radiated power to minimize interference between users.

Users of shared radio frequencies demand some type of access controls so they can avoid interference. Fortunately, as the demands on radio bands have increased, so has the ability to economically use higher frequencies. Expansion to higher frequencies has enabled higher rates of information exchange. But this often results in messages of shorter length, and usually requires sacrificing range or distance between sender and receiver. Higher frequencies are usually limited to line of sight between transmitter and receiver. Most of the new methods for sharing radio frequencies depend on packet radio technology, which is suitable only for digital data transmissions. In one such packet radio technology, wireless LAN, many users may share the same frequency through the use of spread-spectrum technology.

GSM is a wireless telephony standard that is used in most of the world. In the United States GSM shares a pair of frequency bands by using time division and frequency division multiplexing. CDMA is an alternative cellular telephone

radio technology used only in the United States and a few other countries. Advocates of CDMA claim it to be the wireless telephone technology of the future, since it depends on packet switching technology to share the bandwidth.

Packet switching allows the wireless service provider to broadcast variable length assemblies (packets) of digital data continuously from the cell tower to the handset depending on the address content of the packet to be recognized by the handset. In other words, the data broadcast is “packed” with data. The protocol is much like Ethernet, depending upon the receiver to extract the data packets intended for the local terminal. Digital voice data is rarely continuous; no data is sent during pauses in conversation. A second channel at a different frequency is used for the return signal when needed. This makes CDMA very efficient in its use of the spectrum.

GSM and its older form, TDMA, allocate unique time slots for data to be sent from the cell tower to a local handset. If all time slots are being used by handsets, then the call must be routed to another cell tower if it is in range of the receiver. The receiver retains an assigned time slot unless it moves out of range of a cell tower. Each active receiver is assigned a time slot on the receiving frequency, and one on the transmitting frequency. There are many slots at each base frequency.

1.3.5 Privacy/Encryption

Once a radio signal enters the air (or ether, as it is sometimes called), anyone with the necessary equipment may receive it. Wired communications require a physical electrical connection, or at least an inductive coupling that is very close to the wire so as to intercept the signal. Governments have declared that intercepting a wired communication signal is illegal wiretapping and may only be permitted with a court order. No such limitations exist for most types of radio signals. If you transmit, anyone can receive. However, the law in the United States has made listening to some radio signals illegal, even though that is difficult to enforce.

Solutions exist for making radio signals more private. Although no way exists to provide exactly the same level of privacy of an ordinary wired communication, many methods are available for making radio transmissions difficult to interpret, even if we cannot make them impossible to receive. One of the most common ways to achieve privacy is to use highly directional radio antennas, in which

interception is only possible if one has exact knowledge of and access to the line of sight between the sending and receiving antennas. Locating these line-of-sight antennas on towers and rooftops physically limits the potential for interception.

Encryption can make even an intercepted signal difficult or impossible to interpret, hopefully providing privacy equivalent to that of wired communications. Encryption is the science of scrambling the data using a Boolean computational method that depends on a constant called a key. Decryption is the use of a key to unscramble the data and restore it to its original form. An interceptor would need the encryption key to unlock the data and decrypt it, provided that the encryption method is known. Simple encryption is sufficient to protect non-critical or non-vital data, but more complex encryption is required for data exchanges that may involve process, personal, or financial data. Transmissions of data necessary to operate a manufacturing production facility are considered to require high immunity from interception.

There are two types of encryption: secret key and public/private key. Secret key encryption uses a key or cipher consisting of several characters to process the original message using a known method so as to create an encrypted message. The same key is used to decrypt the message after it is received. Many methods, called processes or algorithms, are used for secret key encryption. The best-known algorithm is the Data Encryption Standard (DES). It was developed by the National Institute of Science and Technology (NIST), and is widely published. DES uses a 56-bit secret key. To make it more secure, Triple-DES is sometimes used, in which the same key is processed three times, though the key length is the same. The Advanced Encryption Standard (AES) is the latest NIST development for assuring maximum security of the secret key method. It uses 128-, 192-, and 256-bit keys.

One of the most secure methods for data transmission privacy is the public/private key encryption method, which is used to verify signatures. A user is given a public key that may be published. When the sender “signs” a document, the digital signature is encrypted with the sender’s private key. The encrypted signature and the sender’s public key are both sent to the recipient, who then uses the sender’s public key to verify the signature of the original user.

Document privacy is obtained by encrypting the whole document using the *recipient’s* public key. When the transmission is received, the targeted recipient, and only that recipient, may decrypt the document using his or her own private

key. While complicated, no method provides greater assurance of privacy than public/private key encryption. For public/private key systems to work effectively there must be an open repository for public keys, such as <http://keyserver.pgp.com>, which only support PGP (“Pretty Good Privacy”) encryption keys.

There are two dominant public/private key encryption methods: RSA (Rivest-Shamir-Adleman) and PGP (Pretty Good Privacy). RSA is a product of RSA Security, a company that specializes in security issues. PGP is an open algorithm supported by software from PGP Corporation. Both methods can be used, but PGP is more often used to encrypt an entire message, while RSA is most often used to encrypt signatures and passwords used to log into networks.

Secure socket layer (SSL) is the leading security protocol on the Internet and uses RSA encryption. When an SSL session is started, the server sends its public key to the browser. The browser then uses the public key to send a randomly generated secret key back to the server in order to have a secret key exchange for that session. The problem is that the public key infrastructure (PKI) requires a lot of computational logic. The use of encryption is usually limited to verifying digital signatures and for financial transactions such as a credit card or bank account number, but 128-bit AES encryption has been incorporated into all IEEE 802.15.4 chips, allowing encryption of all transmissions.

1.3.6 Network Membership

Membership in a wired network is achieved by establishing a physical connection to the wiring or to a network element such as a wiring hub or switch. Wireless units are neither connected to nor disconnected from a network. In order to communicate, they must first seek to *join* the wireless network. As part of the protocol for joining the network, a network address is assigned to each device seeking to join the network.

Network membership (joining) is actually a function of the routing capability, which is embedded into all IEEE 802-based networks by using an IEEE 802.1d protocol implemented by network switches (wired) or access points (wireless). The algorithm is called a spanning tree bridge. The network switch or access point learns the address of each connected station when a message is sent from that station, since the FROM address is located in the message header. In this way, messages not intended for the network members for that switch or access

point do not clutter the network. For a station to join the access point's or the switch's local membership list, it must only send a message.

For many years, industry disregarded most network security, but depended on security implemented by servers. With the advent of open access to the Internet, it became necessary to protect local assets by preventing unauthorized access to local area networks. The most popular method used to restrict access to a local area network from the Internet is the VLAN, or Virtual Local Area Network. The LAN has a firewall to prevent direct access to the LAN from the Internet without first being authenticated by the firewall using a VLAN protocol.

Roaming is an essential property of wireless networks, although the need for roaming exists anytime a portable computer is used on different network segments. Any wireless device may physically be moved so as to be in the range of different wireless networks. The ability to roam means that applications should continue to perform their network communications as the device is moved from one wireless network (domain) to another. Networks that support roaming transfer membership transparently from one domain to another.

For a wireless telephony network, roaming is transparent as cell phones move from the range of one cell tower to the next. It's not that simple for mobile computers on a wireless LAN, however. All the wireless LAN access points are usually connected to ports on a single network switch, which performs the routing function. However, doing this results in a clutter of messages being sent to all access points in hopes of finding the targeted station, if it is not on the local membership list of the network switch.

The newest solution for roaming wireless LAN stations is the so-called wireless switch. This is an access point that has the ability to perform advanced 802.1d spanning tree bridging logic. Just like a wired switch, it learns that a station is within its range when the station transmits a message. The problem is that the station may have previously been in the range of a different access point. Recent advances to the IEEE 802.1d standard provide the network management capability to rapidly move the registration of a station from one switch to another. The wireless switch uses this capability to move the registration of a station from one access point/wireless switch to another. By using wireless switches, the broadcast network clutter is reduced.

1.4 Antenna Technology

Although an antenna is usually a passive (non-electronic) element of the wireless network, it is critically important. The antenna(s) on the transmitter couple the signal to the antenna(s) on the receiver just as surely as wire connects wired networks. Furthermore, just as wired network capacity tends to relate to wire size, wireless network throughput depends strongly on antenna *gain*.

One important characteristic of antennas for wireless is that they tend to polarize the transmitted signal. A vertical transmitting antenna will cause a vertical polarization, while a horizontal antenna will cause a horizontal polarization. AM radio is vertically polarized, which is why automobile antennas are vertical. FM broadcast radio is actually horizontally polarized much like VHF and UHF television. Using the same vertical antenna used for AM radio to receive FM radio is a sub-optimal method. However, it still works because polarization is never purely horizontal or vertical, and the antennas are also usually not exactly horizontal or vertical either. Antennas that are not absolutely vertical or horizontal transmit signals that have both horizontal and vertical polarization components. Likewise, receiving antennas that are not absolutely horizontal or vertical can receive signals that have both horizontal and vertical components. Once transmitted signals are reflected during transmission, they may no longer remain purely horizontal or vertically polarized.

1.4.1 Antenna Size

As the frequency of the radio band in use for wireless networks has increased, antennas have become shorter. At the common 2.4-GHz band used for wireless LANs, a full wavelength antenna is only about 12.5 cm (5 inches) long. At these sizes, it becomes possible to integrate antennas entirely inside a product such as an instrument or a notebook computer. When citizens band radio at 27 MHz was popular, base station antennas were over 11m in length. Even automobile antennas at one-fourth wavelength were almost 3m (9 feet) long. The formula for calculating full-wavelength antenna length is the following:

$$\text{wavelength (m)} = 299,792,458 \div \text{frequency (Hz)}$$

The length of the antenna affects the ability to capture the transmitted radio signal. The optimal antenna length is exactly one wavelength. Sub-multiples of the

full wavelength such as one-fourth or one-half wavelength produce less antenna gain, but are also acceptable in most applications with a more convenient size.

1.4.2 Omnidirectional

Most antennas for wireless networks are omnidirectional—they radiate the signal in all directions at the same time. Omnidirectional antennas are the base case and are considered to have zero gain measured in dB. Omnidirectional antennas are usually polarized vertically for convenience.

Since the transmitted energy from an omnidirectional antenna radiates equally in all directions, the signal effectively loses power proportionally to the square of the distance traveled. Additionally, as radio waves pass through any matter they become attenuated in proportion to the density of that matter. In particular, conductive materials such as copper, aluminum, iron, and steel tend to conduct radio energy toward a grounding point, if one exists. Finally, radio waves are received at the targeted radio's antenna; that same antenna will also receive radio energy from other transmitters and even the original signal reflected from an obstruction not in the direct path.

1.4.3 High-Gain Directional

To overcome the signal losses of omnidirectional antennas, directional transmitting antennas concentrate the radiated energy into a narrow beam. Directional receiving antennas capture signals in the near vicinity of the primary receiving antenna and reflect that energy to the primary antenna, thus effectively increasing the signal level or *gain* of the receiving antenna. Several methods are available for concentrating the radio energy, all of which involve reflecting energy emitted in the wrong direction and redirecting it to the target direction. Directional antennas are intended to be aimed manually so the receiving antenna is receiving in the direction of the transmitting antenna. If either the transmitter or receiver is in motion, the antennas must be continuously repositioned to align them.

Directional antennas also tend to eliminate noise, stray signals on the same frequency, and reflections, all of which improves the signal-to-noise ratio, thus improving reception. Directional antennas used for ultra-high frequency (UHF) and microwave radio are illustrated in Figures 2 and 3. Note that with unidirectional service it is not necessary to use a high-gain antenna to both transmit and

receive, nor is it necessary to use the same *type* of high-gain antenna to both transmit and receive. Obviously, to realize the benefits of high-gain antennas in bidirectional service, such as a WLAN, high-gain antennas should be used at both ends to improve service.

The stacked Yagi antenna shown in Figure 2 illustrates a vertically polarized series of antennas. Each element tends to radiate in all horizontal directions, but not vertically. The longer vertical elements behind those in the front are designed to reflect the radiation that is toward the rear, back toward the front. On reception, the rear elements tend to reflect received signals toward the active front antenna elements. This design is generally thought to increase the gain by 3 to 13 dB, depending on the number of vertical elements.

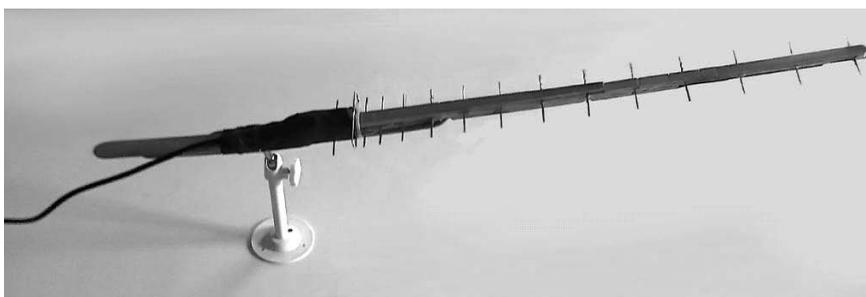


Figure 2. Stacked Yagi Antenna

A parabolic dish, shown in Figure 3, is much more expensive than an omnidirectional or stacked Yagi antenna, but it provides far more gain. The parabolic dish reflects the radiated wave into a narrow beam, and likewise focuses the received energy from a wider area into the receiving antenna. In addition to the higher cost of the antenna and its vulnerability to wind and snow, the parabolic dish suffers from another drawback: the narrow beam width when it is used as a transmitting antenna. The narrow beam width makes the task of aiming the receiving antenna much more difficult, especially when it is used with a parabolic transmitting antenna. A parabolic dish is generally considered to have a gain of 20 to 24 dB over an omnidirectional antenna.



Figure 3. Parabolic Dish Antenna

1.4.4 Planar

The planar antenna design evolved from use in mobile telephones and has become available for commercial use. Planar antennas are small and lightweight, and at wireless LAN frequencies can be embedded into the equipment. Antenna gain can be obtained by building more than one planar antenna into a device. In general, the planar antenna is omnidirectional.

1.4.5 Phased-Array

A phased-array antenna is a two-dimensional organization of planar antennas. Military radar systems were the first to use phased-array antennas. The appeal of the phased-array antenna is that it can exhibit the high gain of a directional antenna and can be aimed electronically without moving the base antenna. Therefore, phased-array antennas enjoy a true advantage when connecting wireless radios in mobile equipment as they move beyond the range of omnidirec-

tional antennas. A simple phased-array antenna is specified by IEEE 802.11ac to enhance data rate by “beam-forming.” As IEEE 802.11ac becomes popular, the beam-forming capability using simple phased-array antennas will become more economical and popular.

Phased-array antennas form a beam electronically rather than by using the reflective properties of metals. Each component planar antenna must be separately driven, with the same signal modified in phase and amplitude to form this beam. Military phased-array antennas use hundreds of elements, but when this technology becomes commercial, many fewer elements will be necessary to achieve a formed beam that is sufficient for industrial distances. Beam formation can usually be directed to an included angle of about 75 degrees, sufficient to help the person installing the router to focus the antennas so that they do not need to point directly at the receiving antenna.

1.5 Wireless Network Topologies

Wired networks have a layout or topology that is determined by the relative physical location of the nodes and network components. Wireless networks are not so easily described. The topology of a wireless network is determined by the logical capabilities of the network components. The user must often determine how the wireless network’s topology is to be configured after installation, or perhaps after some operational experience.

1.5.1 Star

The most typical or default arrangement for a wireless network is a star cluster, in which the wireless access point is at the center, as illustrated in Figure 4. Each wireless device then communicates only with the common access point, which is usually connected via wires to a network switch. Note that there may be more than one access point in the same geographical area, but the assignment of a network node to an access point is logical, not physical. This arrangement then places all of the wireless devices assigned to that access point within the cluster into the same collision domain (meaning messages may overlap) unless there is some means of access control in the protocol on the timing at which a message can be sent.

With Wi-Fi networks there is no protocol for network access control other than to listen before transmitting. If two nodes begin transmitting at the same time on

the same channel, each hearing network silence, both messages will be lost in the resulting collision, since the access point itself will be unable to understand the garbled signal enough to identify the sender, and therefore cannot acknowledge either message. Unfortunately, neither transmitting node will be notified that its message was lost, except when a timeout timer expires after waiting for acknowledgement. When the timeout occurs, both transmitting Wi-Fi nodes will simply try to send their messages again. No defined network protocol exists for Wi-Fi messages to recover from message collision as there is for Ethernet. However, in lightly loaded Wi-Fi networks, message collision is not common.

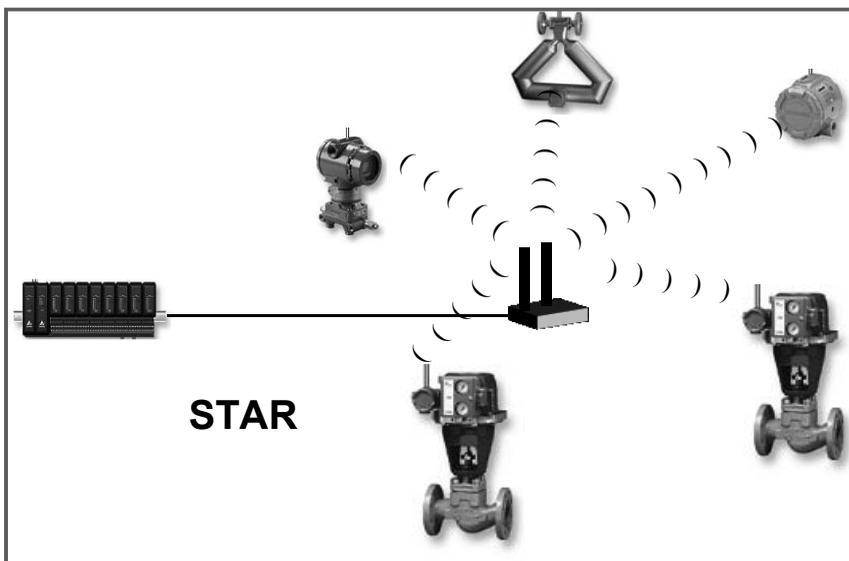


Figure 4. Wireless Star Topology

Wireless access point switches are now appearing for commercial and industrial networks. Their function is similar to that of an unswitched access point, except that they carry a full Layer 2 switching function using the spanning tree bridge protocol, IEEE 802.1d and IEEE 802.1w, and a rapid reconfiguration protocol that is needed for wireless roaming. The spanning tree bridge protocol allows a network switch to learn the address of its connected devices by listening to messages they send. It then routes any messages received at the switch to the device, and to no other. The roaming extension allows a network-connected station to retain network sign-in while moving from the radio zone of one wireless network switch to another. Under 802.1d, when a mobile station moved out of the range

of one switch/access point, a time-out period was necessary before that station could log into another switch/access point. With 802.11w, the station may log into the network at any switch/access point by just sending a message, which will cause it to be logged out of the previous switch/access point without needing the timeout period.

The significance of rapid reconfiguration for industrial automation is obvious in the case of mobile devices such as automated guided vehicles. However, rapid reconfiguration can also be used to increase the reliability of star networks through redundancy to provide the highly reliable networks needed for the manufacturing environment.

When applied to stationary equipment, a wireless network connection is normally highly reliable. Due to interference in the radio spectrum, however, it is possible that messages will not reach the desired destination. In that case, a second switch/access point can provide the redundancy needed for the alternate path required for a highly reliable network. However, in the case of wireless networks, full 100 percent redundancy is not required. Only a viable alternate path that can serve many primary paths can function as a backup path. The alternate path may serve to backup several star clusters.

1.5.2 Tree

As in wired networks, wireless networks can be organized into a tree topology. Each field unit is configured to a network that is connected to a specific switch/access point. That access point is then hierarchically connected to another access point closer to the wired network. The topology appears as illustrated in Figure 5. Notice that the tree network is configured by dedicating channels of communications between access points. In Wi-Fi networks, these dedicated channels may operate at a different frequency from the field devices to access point channels.

1.5.3 Mesh

The newest and most revolutionary form of network is called a mesh. In a mesh network each station is both an end device and a network forwarding element. Mesh networks are naturally self-healing and redundant—exactly the properties needed for industrial automation networks.

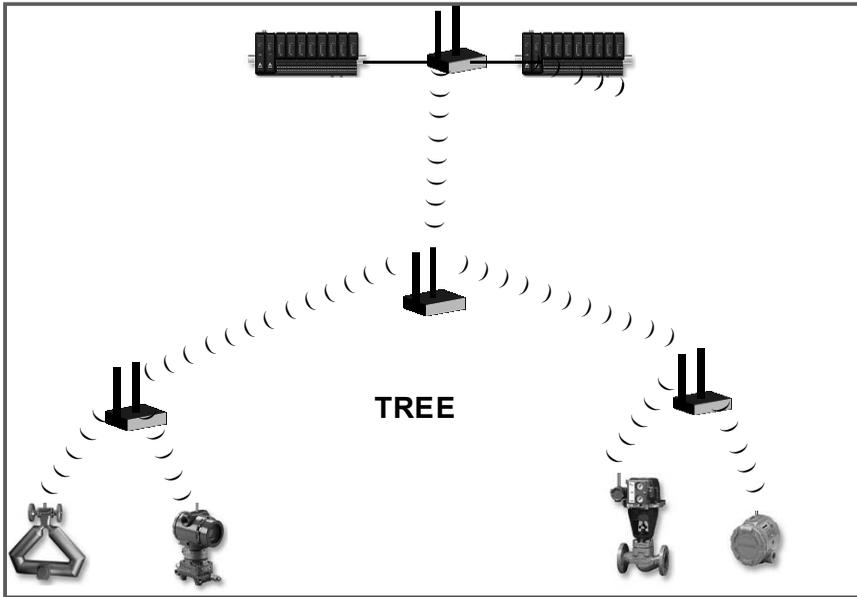


Figure 5. Wireless Tree Topology

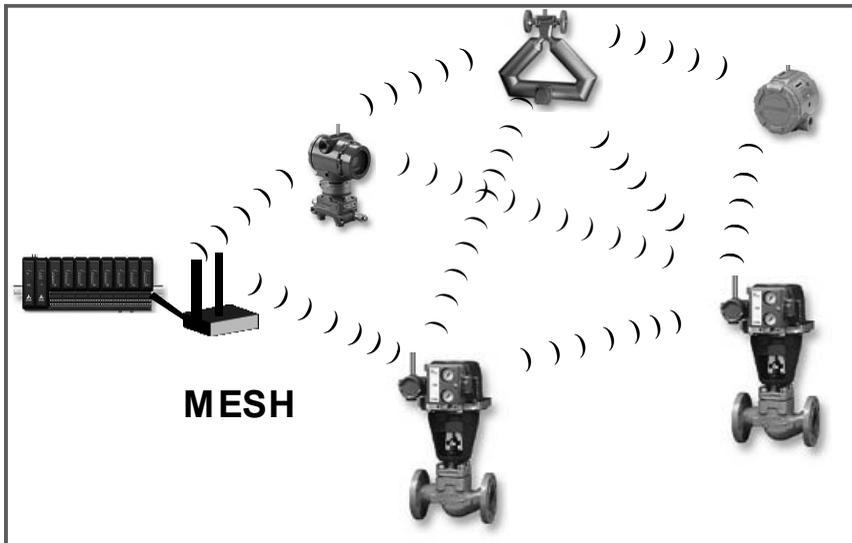


Figure 6. Wireless Mesh Network

In a mesh network, each station is responsible for forwarding a network transmission not intended for itself to other stations within its radio range. Each of those stations, in turn, sends the transmission to at least one other station within

its radio range, as illustrated in Figure 6. Therefore, the network becomes highly redundant, fault-tolerant, and extended in range. The drawbacks are a) that each station must remove duplicate messages, and b) that each hop through the mesh takes time, increasing message latency. In effect, each mesh network station becomes a network router. Additionally, since multiple paths are involved, each receiving station must reject duplicate messages received from divergent paths.

Standardized mesh network protocols also include the capability to build and maintain routing tables so as to provide clues for forwarding messages. This prevents messages from looping in directions other than toward their intended destination, which results in greater network efficiency. Routing tables are dynamically constructed as messages pass through each routing node of the mesh network. Since mesh networks segments that are intended for industrial automation tend to have 256 or fewer nodes, routing tables can be small and the routing simple. Routing tables are updated when new nodes appear in the mesh or when for any reason nodes fail to respond to forwarded messages.

Mesh networks are not new. The Internet itself is a very large wired mesh network with very complex routing algorithms. Since IP addresses do not imply anything about location, messages routed on the Internet hop from one node to another that is (by design) closer to the intended destination. Internet routing algorithms are typically efficient enough that few messages need more than fifteen hops to reach their intended destination.

Wireless mesh networks pose a problem that is not encountered with wired mesh networks such as the Internet. With wireless mesh networks, there is no way, other than by using a highly directional antenna, to prevent a message transmitted by one node of a wireless mesh network from being received by other nodes. This leads to multipath routing, or message duplication. The message identification field of the IP frame is used to identify duplicate messages, which may be discarded if previously received. Multipath routing may also improve network reliability by providing redundant message paths. Using the duplicate message as a means to increase reliability is desired for industrial wireless networks.

The increased latency caused by routing in mesh networks may cause delays in the delivery of messages. Some messages must be delivered to their destination while the data is still “fresh.” Routing may introduce random delays that can make data stale. Network configuration must be adjusted to avoid routing delays for messages that require minimal delays.