The IEEE-802.11 Wireless LAN protocol

We'll next examine devices implementing the IEEE-802.11 family of wireless networking protocols, and get an appreciation of some of the security challenges.

<table>
<thead>
<tr>
<th>TABLE 1. IEEE 802.11 PHY STANDARDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Release date</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>1997</td>
</tr>
<tr>
<td>1999</td>
</tr>
<tr>
<td>1999</td>
</tr>
<tr>
<td>2003</td>
</tr>
<tr>
<td>2009</td>
</tr>
<tr>
<td>2012 (expected)</td>
</tr>
<tr>
<td>2013 (expected)</td>
</tr>
</tbody>
</table>

Often of interest is not simply the maximum possible transmission rate, but the distance over which WiFi may operate.

We need an understanding of the transmission power, propagation, and signal loss over distance and through objects.

The unit dBm is defined as power ratio in decibel (dB) referenced to one milliwatt (mW). It is an abbreviation for dB with respect to 1 mW and the "m" in dBm stands for milliwatt. dBm is different from dB. dBm represents absolute power, whereas dB is a ratio of two values and is used to represent gain or attenuation. For example, 3 dBm means 2 mW, and 3 dB means a gain of 2. Similarly, -3 dBm means 0.5 mW, whereas -3 dB means attenuation of 2.

A WiFi access point (AP) will typically transmit at up to 100mW, and a receiving device will typically be able to discern arriving signal from noise until -90dBm.

<table>
<thead>
<tr>
<th>dBm</th>
<th>mW</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-10</td>
<td>100</td>
</tr>
<tr>
<td>-20</td>
<td>10</td>
</tr>
<tr>
<td>-30</td>
<td>1</td>
</tr>
<tr>
<td>-40</td>
<td>0.1</td>
</tr>
<tr>
<td>-50</td>
<td>0.001</td>
</tr>
<tr>
<td>-60</td>
<td>0.0001</td>
</tr>
<tr>
<td>-70</td>
<td>0.00001</td>
</tr>
<tr>
<td>-80</td>
<td>0.000001</td>
</tr>
<tr>
<td>-90</td>
<td>0.0000001</td>
</tr>
<tr>
<td>-100</td>
<td>0.00000001</td>
</tr>
</tbody>
</table>

Table 1.
Hidden Node, Crouching-Dragon Exposed Node

As a first guess, it would appear that the standard (wired) Ethernet protocol should also work for wireless networks - simply wait until the medium (airwaves) becomes clear, transmit, and then retransmit if a collision occurs.

However, this simple approach will not work, primarily because not all nodes are within range of each other.

Consider these two typical situations:

- (a) A wishes to send a frame to B, but A cannot 'hear' that B is busy receiving a message from C. If A transmits after detecting an idle medium, a collision may result near B. This is described as the hidden terminal or the hidden node problem. C is hidden from A, but their communications can interfere.

- (b) B wishes to transmit to C, but hears that A is transmitting (possibly to someone to the left of A). B incorrectly concludes that it cannot transmit to C, for fear of causing a collision. This is described as the exposed terminal or the exposed node problem.
802.11 Collision Avoidance

A further problem is that most wireless cards are unable to both transmit and receive at the same time, on the same frequency - they employ half-duplex transmissions.

This means that while collisions do occur, they generally cannot be detected (while transmitting).

Unlike their 802.3 wired counterparts, 802.11 wireless LANs do not employ the Carrier Sense Multiple Access with Collision Detection protocol (CSMA/CD). (pedantically, it is incorrect to call 802.11 as being 'wireless Ethernet', but everyone does).

Instead, 802.11 employs collision avoidance to reduce (but not eliminate) the likelihood of collisions occurring. The algorithm is termed Multiple Access with Collision Avoidance (MACA), or CSMA/CA, in which both physical channel sensing and virtual channel sensing are employed.

The basic idea is that before transmitting data frames, the sender and receiver must first exchange additional control frames before the 'true' data frames. The success or failure of this initial exchange either reserves the medium for communication between A and B, or directs how A, B, and all other listening nodes should act.
802.11 Collision Avoidance, continued
Consider the situation of 4 nodes. You can imagine that we have 4 nodes left-to-right, named C, A, B, D.

- A wishes to communicate with B,
- C can hear only A, and
- D can hear only B.

A observes an idle medium, and initially sends a Request to Send (RTS) frame to B. This frame includes a field indicating how long (in microseconds) the actual data frame will be, i.e. how long the sender wishes to hold the medium.

When B receives the RTS it replies with a Clear to Send (CTS) frame, also carrying the length of the data frame.

Any other node hearing the RTS frame (e.g. C) knows that A is making a request, and should not itself transmit until the indicated length/time has elapsed.

Any node hearing the CTS frame (e.g. D) must be close to the receiver (B) and therefore should also not transmit for the indicated length of time.

Any node that hears the RTS frame, but not the CTS frame, knows that it is not close enough to the receiver to interfere, and so is free to transmit (but must first transmit its own RTS...).
802.11 Collision Avoidance, continued

There are two final considerations -

When the receiver (B) successfully receives a data frame, it must reply with an ACK frame. All other nodes must wait for this ACK frame before they can transmit their own RTS frames.

The additional ACK frames were not defined in the early MACA protocols, but added to the Wireless MACA -> MACAW protocol used today.

What if two nodes simultaneously issue an RTS? If they are not in range, there is little problem, we can have two transmission sequences in the same medium. If the two RTS frames collide, then any receivers will not be able to guess what they were, and no CTS frame will be issued.

When no CTS arrives at the senders, they assume a collision and undergo the standard 802.3 binary exponential backoff algorithm.
Access-Point Association

Although two mobile nodes can communicate directly, the more usual approach is for all communication to be through fixed access-points.

We say that a mobile node associates with a single access-point if all of its communications are via that point, and all such related nodes form a cell.

Communication between nodes in different cells requires two access-points connected via a distribution system.

The mechanism employed by a mobile node to select an access-point is termed active scanning:

1. the mobile client node sends Probe frames,
2. all access-points within range reply (if they have capacity) with a Probe Response frame,
3. the mobile node selects an access-point and sends an Association Request frame, and
4. the access-point responds with an Association Response frame.

An alternative is for an access-point to use passive scanning:

1. the access-point periodically sends Beacon frames advertising its existence and abilities (e.g. supported bandwidths),
2. the mobile node may choose to switch to this new access-point using Disassociation and Reassociation frames.

When a node selects a new access-point, the new access-point is expected to inform the old access-point using the distribution system.
Attacks Against Wireless Networks

At the application and transport layers, there is nothing fundamentally different between Denial of Service (DoS) attacks on wired or wireless networks.

However, there are critical differences in the interaction between the network, data-link, and physical layers that increase the risk of DoS attacks in wireless networks:

- **physical-layer attacks**
  attackers of wireless networks need not be close to the network infrastructure, and leave no physical evidence that something has changed (no damaged cabling).

  In addition, an attacker may develop a device to simply saturate the frequency bands with RF noise. This reduces the signal-to-noise ratio, resulting in the receiver not being able to discern any valid signal, or having so many packet collisions in the medium that the bandwidth is cycled to its lowest value (1Mbps). At the moment, cordless phones, microwave ovens, and Bluetooth communications also compete for the standard IEEE-802.11 2.4GHz spectrum.

- **data-link attacks**
  many commercial access-points employ *antenna diversity* - the use of multiple antennae connected to the same access-point. A typical setup includes an access-point connected to a solid wall, with an antenna on either side of the wall.

  The access-point 'learns' on which side of the wall (i.e. via which antenna) a certain MAC address (mobile node) resides, and transmits using that antenna. The dynamic choice of antenna depends on which antenna receives the strongest signal from a particular MAC address.

  An attacker can programmatically change their mobile's wireless card's MAC address to use the MAC address of their victim, and transmit (loudly) from the other side of the wall.
Attacks Against Wireless Networks, continued

- **data-link attacks, continued**
  It is also simple to configure a laptop computer and its wireless card to masquerade as an access-point. If a victim’s node associates with the access-point with the strongest signal, the attacker can monitor, or simply drop, all traffic from the victim.

  Attacks involving the complete termination of communication stream are often termed *black-hole* attacks; ones which only selectively drop or transmit a victim’s packets are termed *grey-hole* attacks.

- **network-layer attacks**
  once a wireless network has acquired a mobile node, the node may perform many standard DoS attacks at the network layer, such as ICMP floods.

  Similarly, a number of innocent ‘DoS attacks’ can be launched simply through an large file-transfer which may flood the wireless infrastructure - we quickly forget how fast a 100Mbps or 1Gbps switched cabled Ethernet is, compared to the shared (hub-like) 11Mbps (cycling downwards through 9Mbps, 2Mbps, 1Mbps) wireless Ethernet.

  For this (and obvious other) reasons, the network segment hosting the access-point should itself be *isolated* from the cabled network through another firewall, and authenticated DHCP or a VPN further used to constrain traffic.

  In particular, the access-point should not be directly connected to the company’s network, and then to the Internet.
Wireless Network Encryption

The popularity and proliferation of commodity wireless networks, such as wireless Ethernet and (to a lesser extent) Bluetooth have exposed networks to new security challenges.

As wireless is a shared medium, all traffic may be intercepted by a potential intruder.

With the range of modern wireless Ethernet (IEEE 802.11b/g) systems being as much as 250 metres (outdoors), a business's network is no longer constrained to where the cable runs - it now includes the car park and the competitor's building across the street.

In the early 2000s, the popular 'sports' of warchalking and drive-by-networking rose to the challenge of finding publicly-accessible (open) wireless networks.

The original designs of the popular IEEE 802.11b/g standards recognised these challenges, and included both encryption and authentication into the standards as early as 1996. Whenever these are being considered anywhere, three factors emerge:

- The need for privacy - how strong must protocols be in terms of computational-power, time, and money?
- Ease of use - if too invasive or too difficult to configure, security will simply be bypassed, and
- Government regulations - encryption products are regulated by many governments which attempt to limit the strength of exported (and imported) algorithms.

We shall look at two forms of WiFi encryption:

- The Wired Equivalent Privacy (WEP) algorithm
- Wi-Fi Protected Access (WPA, WPA2) and 802.11i - will be discussed later in the unit when introducing digital certificates.
The Wired Equivalent Privacy (WEP) algorithm

The Wired Equivalent Privacy (WEP) algorithm is used to protect wireless communication from eavesdropping and to prevent unauthorized access to a wireless network.

WEP relies on a secret key that is shared between a mobile station (a laptop, smartphone, or PDA with a wireless Ethernet card) and an access-point (a base station). The secret key is used to encrypt packets before transmission, and an integrity check is used to detect frame modification in transit.

WEP was selected to meet the following criteria:

- its use is optional for the correct running of a wireless network,
- must provide "reasonably strong" encryption,
- must be self-synchronizing, as nodes may move in and out of range,
- must be computationally efficient, as the encryption must be performed frequently on 'slow' CPUs with limited power supplies, and
- being within (US) government encryption regulations.

The WEP standard does not discuss how the shared key is generated or distributed.

In practice, most installations use a single key that is shared between all mobile stations and access-points using the same Service Set Identifier (SSID).
WEP Encryption

WEP employs a secret key of either a 40- or 104-bits, which are shared between access-points and mobile devices. These secret keys are concatenated with 24-bit initialization vectors (IV) to form what is casually described as 64-bit or 128-bit encryption - marketed as silver or gold wireless Ethernet cards. The 64- or 128-bit key provides input to a pseudo-random number generator (PRNG) - the RC4 encryption algorithm, known as a stream cipher, producing an infinite pseudo-random keystream.

**WEP encryption:**

![Diagram of WEP encryption process]

To prevent against modification in transit, the common CRC-32 checksum (ICV) is calculated over the original plaintext, and appended to produce the total payload. The sender XORs the keystream with the payload to produce ciphertext. This results in ciphertext that is the same length as the original payload.

**WEP decryption:**

![Diagram of WEP decryption process]

The receiver has a copy of the same key, and uses it to initialize and generate the identical keystream. XORing the keystream with the ciphertext yields the original plaintext and the integrity of the original data may be verified by recalculating the CRC-32.
Problems with WEP Encryption

By 2001, a number of research papers and software focused on:

- Passive attacks to decrypt traffic based on statistical analysis.
- Active attacks to inject new traffic from unauthorized mobile stations, based on known plaintext.
- Active attacks to decrypt traffic, based on confusing the access-point.
- Dictionary-building attacks, requiring analysis of about a day's worth of traffic.

The use of RC4 (initially a secret cipher of Ron Rivest, 1987, with a period greater than $10^{100}$, and about 10x faster than DES) is vulnerable to several attacks:

- If an attacker changes a single bit in the ciphertext, the corresponding single bit in the decrypted plaintext will be changed.
- If an eavesdropper intercepts two ciphertexts encrypted with the same keystream, we can obtain the XOR of the two plaintexts. Knowledge of this XOR enables statistical attacks to recover the plaintexts.

The attacks become increasingly practical as more ciphertexts are found. Once one of the plaintexts becomes known, it's trivial to recover all others.
Problems with WEP’s Initialization Vector

To avoid encrypting two ciphertexts with the same keystream, the 24-bit Initialization Vector (IV) is supposed to augment the shared secret key and produce a different RC4 key for each packet. The IV is then included in the packet.

However, both of these measures are frequently implemented incorrectly, resulting in poor security:

- The IEEE 802.11b/g standard does not specify how often the IV should be changed, nor how. As the IV used to encrypt a packet is included in a packet, the receiver does not need to be informed of (nor agree to) any change.

  Most wireless network cards simply initialize the IV to zero when the Ethernet device is initialized (e.g. the PCMCIA card is inserted), and then increments the IV for each subsequent packet. Two cards inserted at roughly the same time provide an abundance of IV collisions for an attacker.

- In addition, the short 24-bit length almost guarantees the reuse of the same keystream under standard conditions.

  A busy access-point, which constantly sends 1500 byte packets at 11Mbps, will exhaust the space of IVs after about 5 hours (even quicker for shorter frames)!

  This allows an attacker to collect two ciphertexts that are encrypted with the same keystream and perform statistical attacks to recover the plaintext. Worse, when the same key is used by all mobile stations in the same Service Set Identifier (SSID), there are even more chances of IV collision.
WEP Authentication

WEP employs the same shared key to authenticate a mobile station as it uses for encryption/decryption. There are two forms of IEEE 802.11b/g authentication:

- **open system authentication** - no authentication, a mobile node may associate itself with any (the closest) access-point and listen to all data sent in plaintext, and
- **shared key authentication** - a shared key is required between a mobile node and an access-point. Again, the IEEE 802.11b/g standard does not specify how the shared access key is distributed amongst mobile nodes and, to support roaming, a number of access-points in a cluster (e.g. in the UWA-UNIfly network) will all have the same key.

The authentication sequence is a challenge-response scheme, as follows:

1. a roaming mobile node sends an `AuthenticationFrame` to the access-point (AP),
2. the AP replies with an `AuthenticationChallengeText` frame consisting of 128 random bytes (the challenge),
3. the mobile node encrypts the bytes with (what it hopes is) the shared key and returns them to the AP, and
4. the AP decrypts the returned text, and compares it to the original challenge.

Once successful, the AP will accept packets from the mobile node’s MAC address.
Attacks by Patient Attackers

**Passive traffic decryption** -
A passive eavesdropper can intercept all wireless traffic, until an IV collision occurs. By XORing two packets that use the same IV, the attacker obtains the XOR of the two plaintext messages.

The resulting XOR can be used to infer data about the contents of the two messages. IP traffic is often very predictable and includes a lot of redundancy.

An extension to this attack uses a host somewhere on the Internet to send traffic from the outside to a host on the wireless network. The contents will be known to the attacker, yielding known plaintext. The attacker intercepts the encrypted version of the messages, and be able to decrypt all packets using the same IV.

**Active traffic injection** -
an attacker knowing the exact plaintext for one encrypted message can use this knowledge to construct correct encrypted packets.

This is achieved by constructing a new message, calculating the CRC-32, and performing bit flips on the original encrypted message to change the plaintext to the new message.

**Table-based attacks** -
The small space of possible initialization vectors allows an attacker to build a decryption table (size 15GB). Once the plaintext for a known packet is known the attacker computes the RC4 key stream generated by the IV used.

This key stream can be used to decrypt all other packets using the same IV. Over time, the attacker builds a table of IVs and corresponding key streams. The attacker can then decrypt every packet.
Addressing WEP problems

The many attacks on WEP left WiFi networks without viable link-layer security.

The Temporal Key Integrity Protocol (TKIP) was designed in 2002, as part of the IEEE 802.11i standard, to replace WEP without requiring the replacement of legacy hardware.

TKIP is often described as a wrapper around WEP encryption. TKIP employs the same encryption ideas and RC4 algorithm defined for WEP. However, TKIP's encryption key is 128 bits long - addressing WEP's first problem of WEP - its short key length.

TKIP changes the key used for each packet. The key is created by mixing a base key (the Pseudo Transient Key), the MAC address of the transmitting node, and the sequence number of the packet.

Each TKIP packet has a unique 48-bit sequence number, incremented for each new packet, and used both as the IV and part of the key.

Placing a sequence number into the key ensures that the key is also different for every packet. This addresses WEP's next problem - avoiding "collision attacks", which occur when the same key is used for two different packets.

Having the 48-bit sequence number of the packet be the IV addresses another WEP problem - "replay attacks". Old, replayed packets will be detected as out of order.