WHITE PAPER

Industrial IoT in a 5G world - Technology

aruba
a Hewlett Packard
Enterprise company

ABSTRACT

An Industrial IoT system is a complex architecture encompassing sensors, communications, big-data storage, edge computing and advanced analytics among its disciplines. In the communications segment, the 'last hop' for Industrial IoT is now wireless wherever possible, driving cost saving, increased flexibility and greater mobility than wired connections. When considering the last hop, a key question, which we will investigate in this paper, is which wireless technology to select: for large-scale Industrial IoT networks, the viable technologies are variants of the Wi-Fi used in enterprise networking today, and 4G from public or private cellular networks.

The paper examines the underlying technical strengths and weaknesses of each technology in the network models used today, comparing Wi-Fi 5 (also known as 802.11ac) with public and private 4G architectures. It also covers the emerging Wi-Fi 6 (also known as 802.11ax) and 5G technologies which will become available to the Industrial IoT market in the near future.

Technology comparison is a very complex exercise, where assumptions and models exert a huge influence on the results. Network architects looking for evidence on which to base a decision would be well-advised to wait until equipment is available and to test for themselves. But in the interim, we outline in this paper the issues involved in such a comparison.

EXECUTIVE SUMMARY

When setting out to compare the capabilities of two rapidly-evolving standards, it is first necessary to decide which technologies to compare. In this paper we concentrate on Wi-Fi 5 and 4G, LTE Advanced, as defined in 3GPP r13: these are the technologies embodied in the iPhone 7 Plus and similar devices. Even then the comparison is difficult, as commercial devices do not implement every feature in the standards. Therefore, we set out to pick equivalent state-of-the-art implementations at a point in time.

The first section investigates spectral efficiency, how many bits/second each technology can move in a given RF channel width. This section concludes that at the fundamental level of the OFDM symbol, the two are almost exactly equivalent – there is no practical difference. Since Wi-Fi 6 is rolling out ahead of 5G, Wi-Fi in real networks will enjoy a symbol-level advantage due to its 1024 QAM modulation for some years.

It is at the MAC layer that divergences occur, due to the very different implementations. 4G, in FDD form in licensed spectrum, uses a framed, OFDMA format where nodes are assigned varying numbers of Resource Blocks in frequency and time within a fixed structure. This translates to a defined amount of resource to send traffic and, given a target error-rate with a known link budget, the required bit-rate can be computed. In this OFDMA structure there is little ambiguity about bit-rate: it is made possible because there is one controlling node in the network, the base station that schedules all transmissions, both downlink and uplink.

Wi-Fi 5 uses a very different MAC, a CSMA/CA scheme in which nodes make autonomous decisions about when they transmit, it is a loosely-coupled system with no central control. Any node wishing to transmit first senses whether there is a packet on the air. If not, it can transmit immediately. But otherwise it sets a random backoff timer and when that expires, it senses again for clear air and repeats. This protocol is well-suited to unlicensed spectrum and uncoordinated nodes, as it can accommodate large numbers of overlapping networks and interference from diverse sources. It also adapts readily to bursty data traffic and changes in uplink-downlink traffic balance. But it is inherently less spectrally efficient than an OFDMA structure; a good deal of airtime can be lost to the contention mechanism, and any airtime not used for packet transmission subtracts from spectral efficiency. We estimate that a straight comparison of FDD 4G against Wi-Fi 5 shows a ~35% advantage in spectral efficiency for 4G.

But this is not necessarily a fair comparison, for a number of reasons which we identify. First, Wi-Fi uses TDD not FDD and this is responsible for around 30% of the difference. 4G is able to use FDD in some licensed bands, but as it moves into 'private LTE' spectrum like CBRS, it needs to use TDD, as will be the case in unlicensed bands. In these cases, the ~35% difference narrows to ~25%.

Next, 4G enjoys sheltered operation in licensed bands, as it is guaranteed freedom from intentional interference in exchange for paying for the license. As the technology is moved into unlicensed bands, and a factory environment with potentially high levels of RF interference from machinery and other sources, error rates and retransmissions will rise. 4G deals with this by using features like HARQ that add forward error-correction, and retransmitting following errors, but this all adds overhead and detracts from spectral efficiency, while the Wi-Fi contention mechanism is already set up to expect these impairments and suffers little efficiency impairment.

Also, Wi-Fi is a data-oriented protocol: it expects the rapidly-changing load profiles that result from bursty data traffic. Since it makes packet-by-packet transmission decisions, it reacts instantly when, for instance, a node starts streaming video or a file transfer begins. But the control mechanisms of 4G OFDMA take time to react by reassigning Resource Blocks as nodes change traffic profiles, and in TDD to changes in the uplink-downlink traffic balance. Similarly, to allow for higher-priority QoS, OFDMA must over-provision Resource Blocks to allow a buffer space for extra traffic – but this empty buffer space counts against spectral efficiency. Wi-Fi 5, where modified backoff parameters give priority for QoS, is less affected by rapidly-changing load profiles.

We conclude that, when comparing practical deployments of 4G technology in Industrial IoT settings with the equivalent Wi-Fi 5 network, the nominal spectral efficiency advantage of 35% enjoyed by the public, licensed cellular network is reduced to less than 10%. And with the new Wi-Fi 6 generation which adds OFDMA, the gap is completely closed.

In any event, spectral efficiency, while useful to some to claim bragging rights, is not a significant differentiator in practical networks. The achievable bit-rates, error rates, range and other parameters are more important, and these are determined by many considerations above spectral efficiency: MIMO, channel width, client capabilities and others. Nevertheless, it provides an opportunity to compare underlying MAC and PHY protocols.

A more detailed discussion of interference differentiates between intentional transmitters, which may be present in unlicensed bands but not in licensed, and other sources of interference. For instance, most 4G networks are configured with a frequency re-use factor of 1, where the same channel is used everywhere. This maximizes network capacity but gives rise to interference where cells overlap – albeit interference that can be managed by the network operator. Most private 4G networks will be allocated only one channel, so they will need to operate in this way. Licensees of the PAL tier of the forthcoming CBRS band may be pre-empted by incumbent users appearing over the horizon. This should only happen near the coastline, but the exclusion zone in such cases stretches many miles inland. This is perhaps the ultimate in interference. And although the bands used by Wi-Fi are unlicensed and available to all, an industrial plant covering a large geography, particularly if it is situated in a rural area, will effectively manage its own airspace, as Wi-Fi transmitters on adjacent property will not be of sufficient power to cause problems in operational areas. Interference is a nuanced effect in Industrial IoT networking.

Meanwhile QoS, and general reliability are important considerations for OT (Operational Technology) applications, where plant availability and manufacturing output targets are paramount considerations. QoS reflects that all traffic flows are important, but some are more important than others, while particular applications may have special reliability requirements. We emphasize here that, while 4G OFDMA has the potential to provide strong QoS, the public network effectively implements only one level of service, and it is surprisingly complicated for operators of the public cellular network to build, price and provision new services – it remains to be seen what they will offer in the way of tiered QoS options for Industrial IoT. Similarly, private 4G networks may have the technology but this needs to be translated to operational services in many cases. While Wi-Fi 5 uses a less deterministic QoS mechanism, its behavior is well understood, and enterprises are used to planning and configuring it – and coordinating QoS in the wireless network with the wired network for end-to-end performance is a natural fit with Wi-Fi. The advent of OFDMA in Wi-Fi 6 enables much stronger Wi-Fi QoS, very similar to 4G/5G capabilities.

When considering the range of the signal, 4G in licensed bands has a significant advantage over Wi-Fi due to higher transmit power and antenna gain by regulation, and lower-frequency operating bands. But some of this advantage disappears in the CBRS mid-band spectrum, and the rest is obviated for unlicensed operation, as the rules there are identical for 4G and Wi-Fi technologies. If high sensor density or traffic capacity per unit area is significant, cells will need to be smaller and the range advantage will disappear. This will accelerate as 5G brings more small cells to public and private cellular networks.

Any IT or OT engineer architecting an Industrial IoT network will consider availability of devices. Wi-Fi is well-known and has been available for many years, building up a huge ecosystem of devices, while cellular technology, tied for so long to cellphones and the public network, is much less developed. Perhaps the best option for 4G is to use local aggregators or gateways, adapting from wired IoT protocols or even from Wi-Fi at the last-hop local network to a 4G connectivity tier.

Mobility is an area where 4G networks have a clear advantage over Wi-Fi. Whereas the latter is effective to ~100 km/hr, 4G and 5G can serve clients moving at greater than 200 km/hr. And where clients are moving on public roads, the public cellular network with its extensive wide-area coverage may be essential to connectivity.

The comparison ends with a discussion of security: identity, authentication and encryption. Wi-Fi enables a broad range of security options, ranging from open networks – no security – through pre-shared keys to full 802.1X authentication with military-grade encryption. The public perception is often colored by the less-secure (but easier-connected) options consumers encounter at home and when shopping, but enterprises have for many years configured their WLANs to the highest levels of security, similar to or exceeding those used in 4G cellular networks.

The most significant differences in security are that Wi-Fi uses IETF standards to allow authentication with a variety of identities: passwords, X.509 certificates, even SIM cards, while the 4G network has a proprietary though widely-used scheme that is still tied to the SIM identity. Any organization building a private 4G network will need to create its own SIM cards. This is set to change with the 5G architecture, but even then, Wi-Fi will offer more flexibility and ease of configuration for security.

There are several more dimensions to a choice of Industrial IoT technology, such as the availability and cost of spectrum, which are not considered here. They are dealt with in a companion paper on Industrial IoT architectures.

We conclude that the underlying technology used in the Wi-Fi and 4G/5G cellular standards is converging, and it is not possible to say that one is significantly superior to the other in Industrial IoT settings. 4G can offer longer range from the base station – but not under all conditions – and, if roaming to the public cellular network, it extends coverage across long distances off-premises. It also works at freeway speeds, where Wi-Fi is limited to ~100 km/hr and is ideal for applications like AGVs (Autonomous Guided Vehicles). But Wi-Fi, in unlicensed spectrum, is universal and completely under the control of the industrial customer in a way that 4G/5G networks may not be, either because of licensing restrictions or equipment and configuration complexity.

The remainder of this paper compares Wi-Fi and cellular technology capabilities in a number of dimensions:

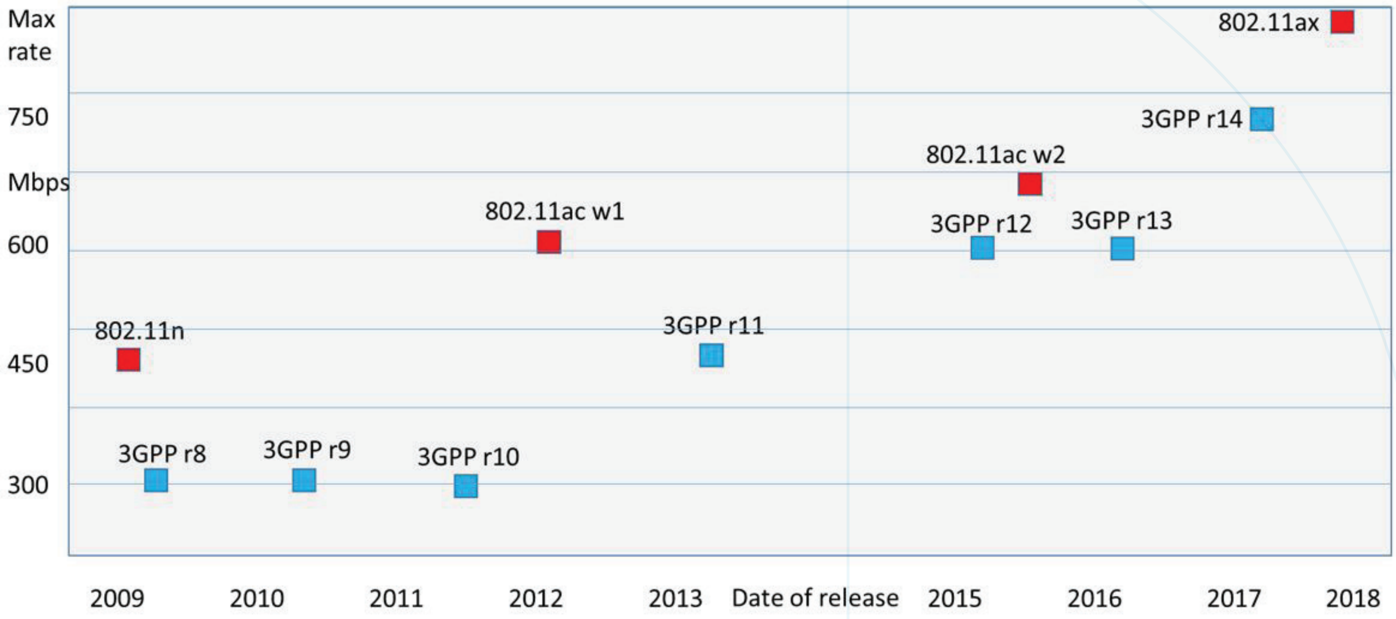
TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TIMING OF STANDARDS REVISIONS

Many technology comparisons go little deeper than "X is faster than Y" bullets on a slide. But these technologies are complex and cannot be summarized without a documented set of assumptions. For the technologies discussed here, which are continually improving, perhaps the first assumption is what point in time, or revision on the standards to take for a comparison.



The chart above graphs data-rate vs date for various 3GPP and Wi-Fi standards, showing major revision levels. It shows equivalent performance, in peak data-rate, over time.

At any time, the Wi-Fi 802.11 and 3GPP standards development organizations have a number of task groups in progress, each working on a different aspect of the standards and delivering updates or amendments as their tasks complete. Wi-Fi tends to produce a new PHY (Physical Layer) standard every 3-5 years, while 3GPP revisions are approximately annual or biennial. (We will only consider the PHY and related MAC amendments that affect data-rates and performance in this section. These standards have many other facets including management, security, location, service discovery to name a few.)

The graph shows that Wi-Fi tends to jump ahead of 3GPP with new amendments, but as its releases are farther apart in time, 3GPP catches up before the next Wi-Fi release. We could add that Wi-Fi is incorporated into new access points and devices very quickly after (or sometimes a little before) release of the standard, whereas 3GPP tends to greater lag times. For instance, many people like to compare 802.11ax (Wi-Fi 6) to 5G (5G-NR new radio in this case) but the Samsung Galaxy S10 became commercially available with Wi-Fi 6 in March 2019, whereas we do not expect the iPhone to adopt 5G-NR until late 2020 or 2021, with similar lag times for network build-outs.

(Even the date of standard releases is debatable. The IEEE has many 'final' revisions and member ballots before a standard is officially released, often 9 months or more after it was practically frozen, and equipment became commercially available.)

The rates shown above are normalized for MIMO (Multiple-Input, Multiple-Output) at a single spatial stream, as adding spatial streams increases data rates proportionally, and this is one area where implementations lag standards. MIMO shows the importance of separating performance theoretically attainable in the standard from what contemporary equipment can achieve. For example, the 802.11n amendment specified up to 4x spatial streams but it was several years before 4x4 APs became available, and most cellphones are still limited to 2x2 for antenna-configuration and cost reasons among others. Which figure to pick, the 'theoretical' measure of what the standard allows, or a 'practical' alternative based on currently-available equipment? We avoid the question by normalizing to 1x1 for comparison purposes.

Similarly, the data rates are normalized to a 10 MHz RF channel. Modern radio systems are capable of operating in very wide channels – 160 MHz for Wi-Fi – or aggregating many channels – 8 or more for 3GPP – to achieve eye-watering theoretical data rates, even though no mainstream equipment or networks support these capabilities. The normalization along both MIMO and channel width allows a more sensible comparison, in our opinion.

For most of this paper we will compare 802.11ac wave 2 with 3GPP r12 or r13, as these are the technologies found together in the iPhone 7 Plus and similar devices.

SPECTRAL EFFICIENCY

The first technical comparison we will make is of spectral efficiency. Spectral efficiency defines how much data, measured in Mbps or Gbps, can be carried in a given amount of spectrum, measured in MHz. Radio experts like to use spectral efficiency as a measure of the sophistication of a technology. It also has a practical use as an indication of the maximum data rates attainable given a certain channel bandwidth. Spectral efficiency can be a useful measure, but it is often gamed by assuming unrealistic measures of channel aggregation, MIMO spatial multiplexing gain and other factors that are not practically attainable by real-world equipment under real-world conditions (see the discussion above on comparison of data-rates).

Nevertheless, as it is a widely-quoted measure where Wi-Fi and 4G/5G technologies are compared, we include a technical assessment below.

Spectral Efficiency at the Physical Layer

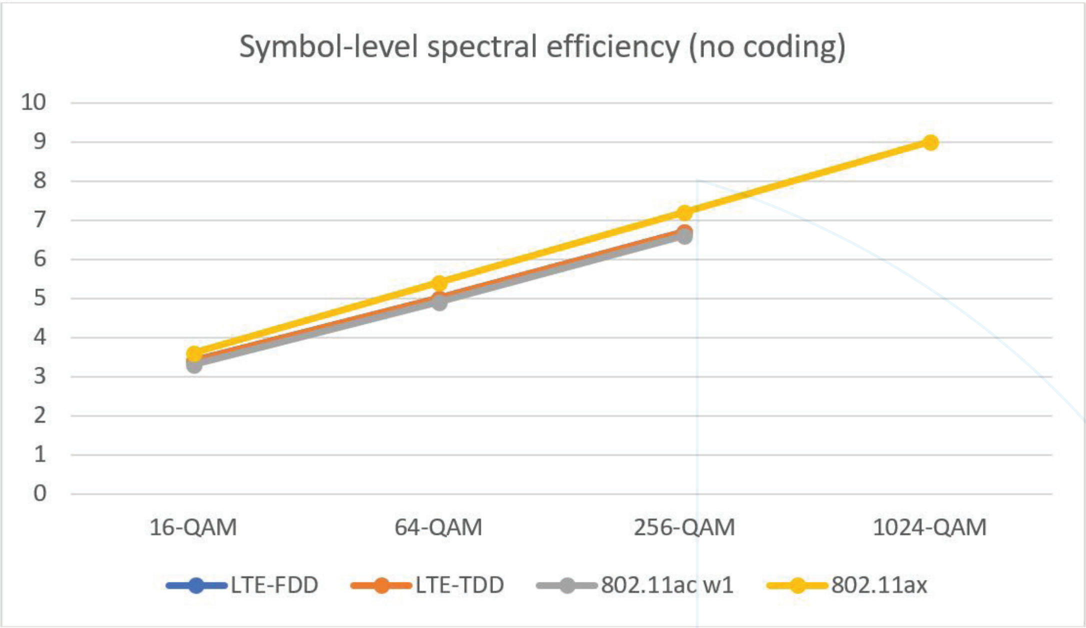
The OFDM (Orthogonal Frequency Division Multiplexing) symbol is the underlying physical-layer unit of transmission for both Wi-Fi and 4G/5G protocols. This calculation considers how many OFDM symbols are transmitted per second (assuming they are transmitted back-to-back without interruption) and how many bits are carried per symbol (there is also a trade-off between bits per symbol and error rate). It gives an approximation to the maximum bit-per-second rate.

Spectral efficiency is calculated by dividing the bps rate by the bandwidth used for transmission, to determine bps/Hz.

The bottom line reveals that the 'symbols/sec/Hz/stream' figures are almost identical. The difference is in the number of bits/symbol, or QAM level. The standards could define arbitrary levels of QAM but chip technology – independent of LTE or Wi-Fi – only allows a certain level at any point in time, so this difference should not be meaningful for our comparison. The analysis below shows how the QAM level determines fundamental spectral efficiency.

(Note again that in the calculations throughout this paper we discount MIMO (Multiple Input, Multiple Output) effects and normalize to 1x spatial stream. MIMO multiplies the capacity of a connection by using multipath effects: two spatial streams have twice the data-carrying capacity of one. But to take advantage of multiple streams, devices must have extra RF components and antennas, one per stream. So different devices support varying numbers of MIMO streams. Since the MIMO effect is implementation-dependent and would swamp all other spectral efficiency factors, all calculations in this paper are for single streams.)

LTE		Wi-Fi	
Channel width MHz	10	Channel width MHz	80
MIMO spatial streams	1	MIMO spatial streams	1
Subframes per second	1000	Symbols per second per subcarrier	280000
Slots per subframe	2	Usable subcarriers	234
Resource blocks per slot	50		
Subcarriers per resource block	12		
Symbols per subcarrier	7		
Bits per symbol (64 QAM)	6	Bits per symbol (256 QAM)	8
Spectral Efficiency (bits/sec/Hz/stream)	5.04	Spectral Efficiency (bits/sec/Hz/stream)	6.55
(comprised of)	0.84 symbols/sec/Hz/stream x 6 bits/symbol	(comprised of)	0.82 symbols/sec/Hz/stream x 8 bits/symbol



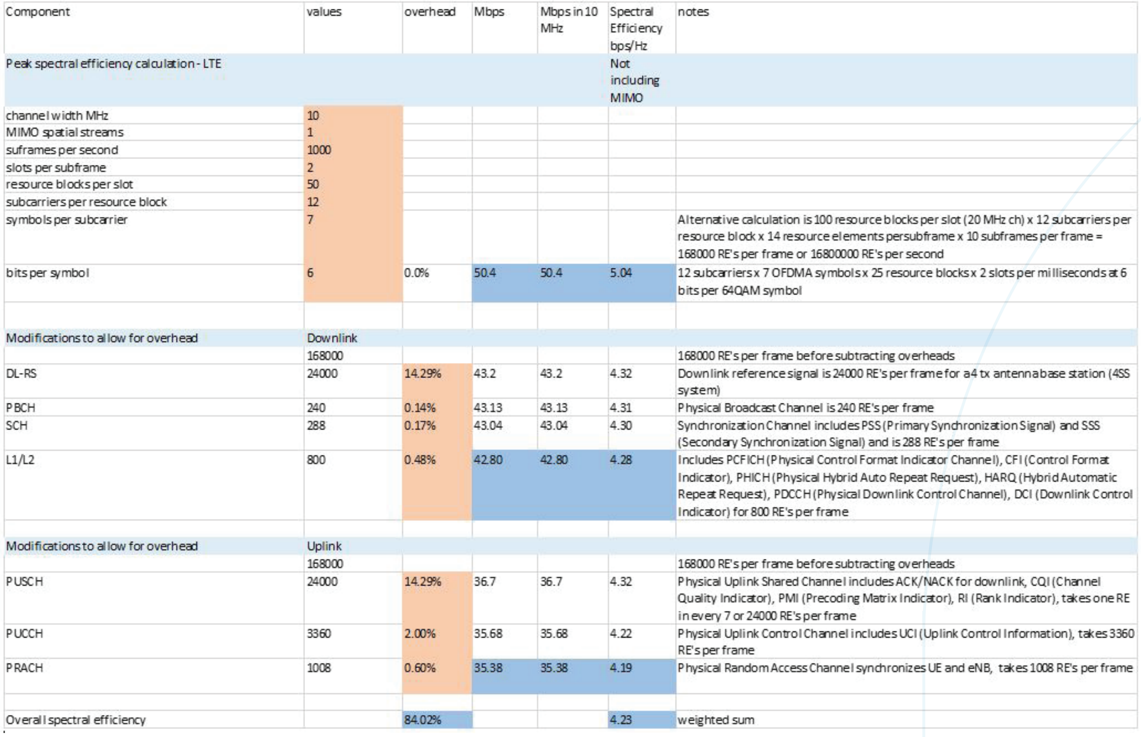
At the physical layer, LTE-FDD and LTE-TDD are the same. In fact, all the LTE and Wi-Fi technologies are so close as to be indistinguishable, once adjusted to the same level of QAM. As will be noted later, LTE implementations in devices tend to lag Wi-Fi in QAM terms, for instance the iPhone 7 Plus has 64 QAM for LTE downlink and 16 QAM uplink, but 256 QAM uplink and downlink for Wi-Fi.

As the title notes, neither the Wi-Fi nor LTE figures above include coding overhead. It is added in the next section.

Spectral Efficiency for the whole system

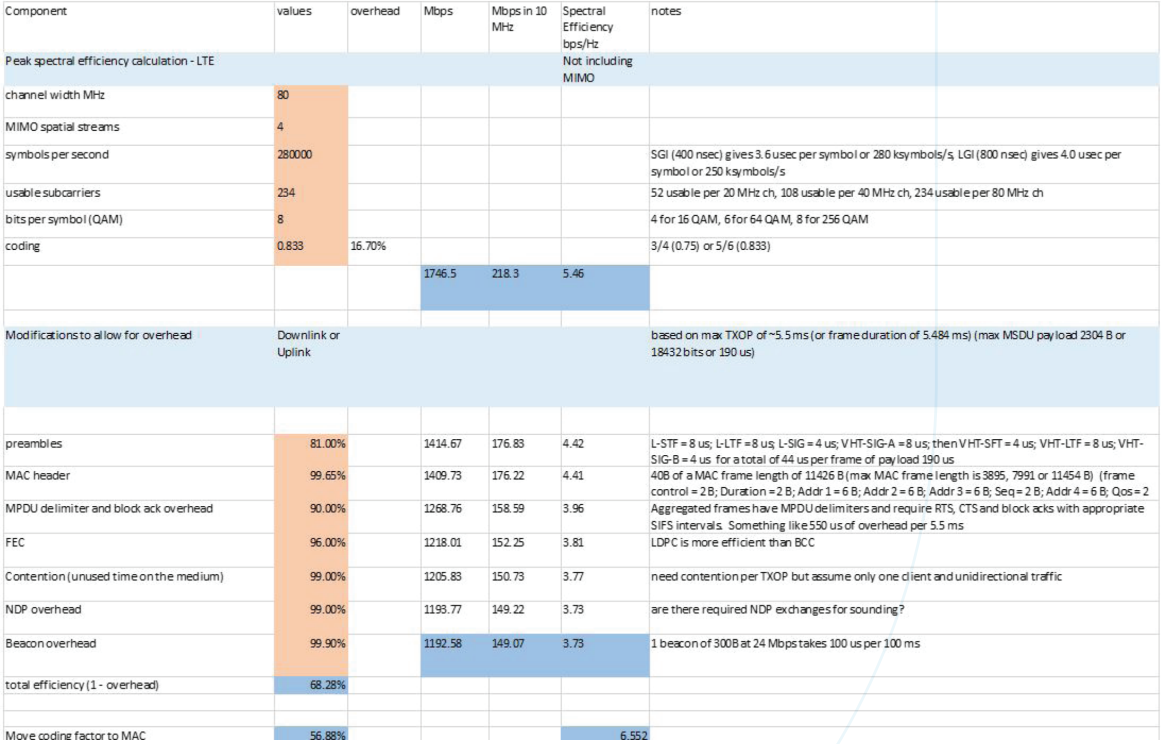
This section takes the physical layer spectral efficiency calculated above and adds in all the factors necessary to make a working link, but taking best-case values.
It is equivalent to the throughput of a single base station / AP with a single UE terminal / client device at short range in a shielded room.

Next is a calculation for LTE-FDD. The top section deals with the physical layer, as discussed above. Below that, the downlink and uplink streams are considered separately, as they have different overheads. The new factors include reference signals and control channels. Note that this is a best-case scenario, with no coding or error correction, and any practical network would not achieve the level of ~4.2 bits/sec/Hz/stream.



Component	values	overhead	Mbps	Mbps in 10 MHz	Spectral Efficiency bps/Hz	notes
Peak spectral efficiency calculation - LTE					Not including MIMO	
channel width MHz	10					
MIMO spatial streams	1					
suframes per second	1000					
slots per subframe	2					
resource blocks per slot	50					
subcarriers per resource block	12					
symbols per subcarrier	7					Alternative calculation is 100 resource blocks per slot (20 MHz ch) x 12 subcarriers per resource block x 14 resource elements persubframe x 10 subframes per frame = 168000 RE's per frame or 16800000 RE's per second
bits per symbol	6	0.0%	50.4	50.4	5.04	12 subcarriers x 7 OFDMA symbols x 25 resource blocks x 2 slots per milliseconds at 6 bits per 64QAM symbol
Modifications to allow for overhead	Downlink					
	168000					168000 RE's per frame before subtracting overheads
DL-RS	24000	14.29%	43.2	43.2	4.32	Downlink reference signal is 24000 RE's per frame for a4 tx antenna base station (4SS system)
PBCH	240	0.14%	43.13	43.13	4.31	Physical Broadcast Channel is 240 RE's per frame
SCH	288	0.17%	43.04	43.04	4.30	Synchronization Channel includes PSS (Primary Synchronization Signal) and SSS (Secondary Synchronization Signal) and is 288 RE's per frame
L1/L2	800	0.48%	42.80	42.80	4.28	Includes PCFICH (Physical Control Format Indicator Channel), CFI (Control Format Indicator), PHICH (Physical Hybrid Auto Repeat Request), HARQ (Hybrid Automatic Repeat Request), PDCCH (Physical Downlink Control Channel), DCI (Downlink Control Indicator) for 800 RE's per frame
Modifications to allow for overhead	Uplink					
	168000					168000 RE's per frame before subtracting overheads
PUSCH	24000	14.29%	36.7	36.7	4.32	Physical Uplink Shared Channel includes ACK/NACK for downlink, CQI (Channel Quality Indicator), PMI (Precoding Matrix Indicator), RI (Rank Indicator), takes one RE in every 7 or 24000 RE's per frame
PUCCH	3360	2.00%	35.68	35.68	4.22	Physical Uplink Control Channel includes UCI (Uplink Control Information), takes 3360 RE's per frame
PRACH	1008	0.60%	35.38	35.38	4.19	Physical Random Access Channel synchronizes UE and eNB, takes 1008 RE's per frame
Overall spectral efficiency		84.02%			4.23	weighted sum

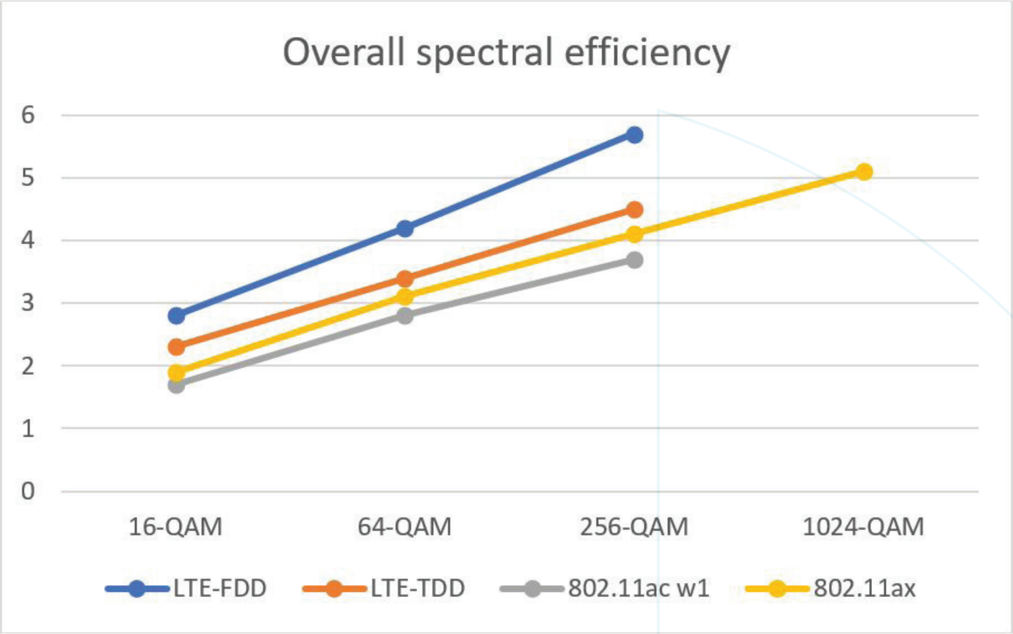
Below is an equivalent calculation for Wi-Fi:



Component	values	overhead	Mbps	Mbps in 10 MHz	Spectral Efficiency bps/Hz	notes
Peak spectral efficiency calculation - LTE					Not including MIMO	
channel width MHz	80					
MIMO spatial streams	4					
symbols per second	280000					SGI (400 nsec) gives 3.6 usec per symbol or 280 ksymbols/s, LGI (800 nsec) gives 4.0 usec per symbol or 250 ksymbols/s
usable subcarriers	234					52 usable per 20 MHz ch, 108 usable per 40 MHz ch, 234 usable per 80 MHz ch
bits per symbol (QAM)	8					4 for 16 QAM, 6 for 64 QAM, 8 for 256 QAM
coding	0.833	16.70%				3/4 (0.75) or 5/6 (0.833)
			1746.5	218.3	5.46	
Modifications to allow for overhead	Downlink or Uplink					based on max TXOP of ~5.5 ms (or frame duration of 5.484 ms) (max MSDU payload 2304 B or 18432 bits or 190 us)
preambles	81.00%		1414.67	176.83	4.42	L-STF = 8 us; L-LTF =8 us; L-SIG = 4 us; VHT-SIG-A =8 us; then VHT-SFT = 4 us; VHT-LTF = 8 us; VHT-SIG-B = 4 us for a total of 44 us per frame of payload 190 us
MAC header	99.65%		1409.73	176.22	4.41	40B of a MAC frame length of 11426 B (max MAC frame length is 3895, 7991 or 11454 B) (frame control = 2 B; Duration = 2 B; Addr 1 = 6 B; Addr 2 = 6 B; Addr 3 = 6 B; Seq = 2 B; Addr 4 = 6 B; Qos = 2
MPDU delimiter and block ack overhead	90.00%		1268.76	158.59	3.96	Aggregated frames have MPDU delimiters and require RTS, CTS and block acks with appropriate SIFS intervals. Something like 550 us of overhead per 5.5 ms
FEC	96.00%		1218.01	152.25	3.81	LDPC is more efficient than BCC
Contention (unused time on the medium)	99.00%		1205.83	150.73	3.77	need contention per TXOP but assume only one client and unidirectional traffic
NDP overhead	99.00%		1193.77	149.22	3.73	are there required NDP exchanges for sounding?
Beacon overhead	99.90%		1192.58	149.07	3.73	1 beacon of 300B at 24 Mbps takes 100 us per 100 ms
total efficiency (1 - overhead)	68.28%					
Move coding factor to MAC	56.88%				6.552	

The connection is symmetrical, so uplink and downlink are identical. Here the coding factor is considered at the physical layer, as is usually the case with Wi-Fi, but the overall result is unchanged by this move. Contributors to overhead include the MAC header, inter-frame spacing, request-to-send and clear-to-send frames and beacons.

A per-QAM comparison at this level:



As one might expect, the FDD mode of LTE is the most efficient, while the TDD mode has a 10% advantage over 802.11ax and 21% over 802.11ac wave 1, for any given level of QAM. As noted below, practical implementations of LTE are usually behind Wi-Fi in terms of QAM level.

4G/5G in the 5 GHz band: LAA and MulteFire deployments, and relative spectral efficiency

The section above dealt with the PHY and MAC layers, but changes need to be made to LTE at higher levels in order to operate in the 5 GHz band. Here are some of the considerations:

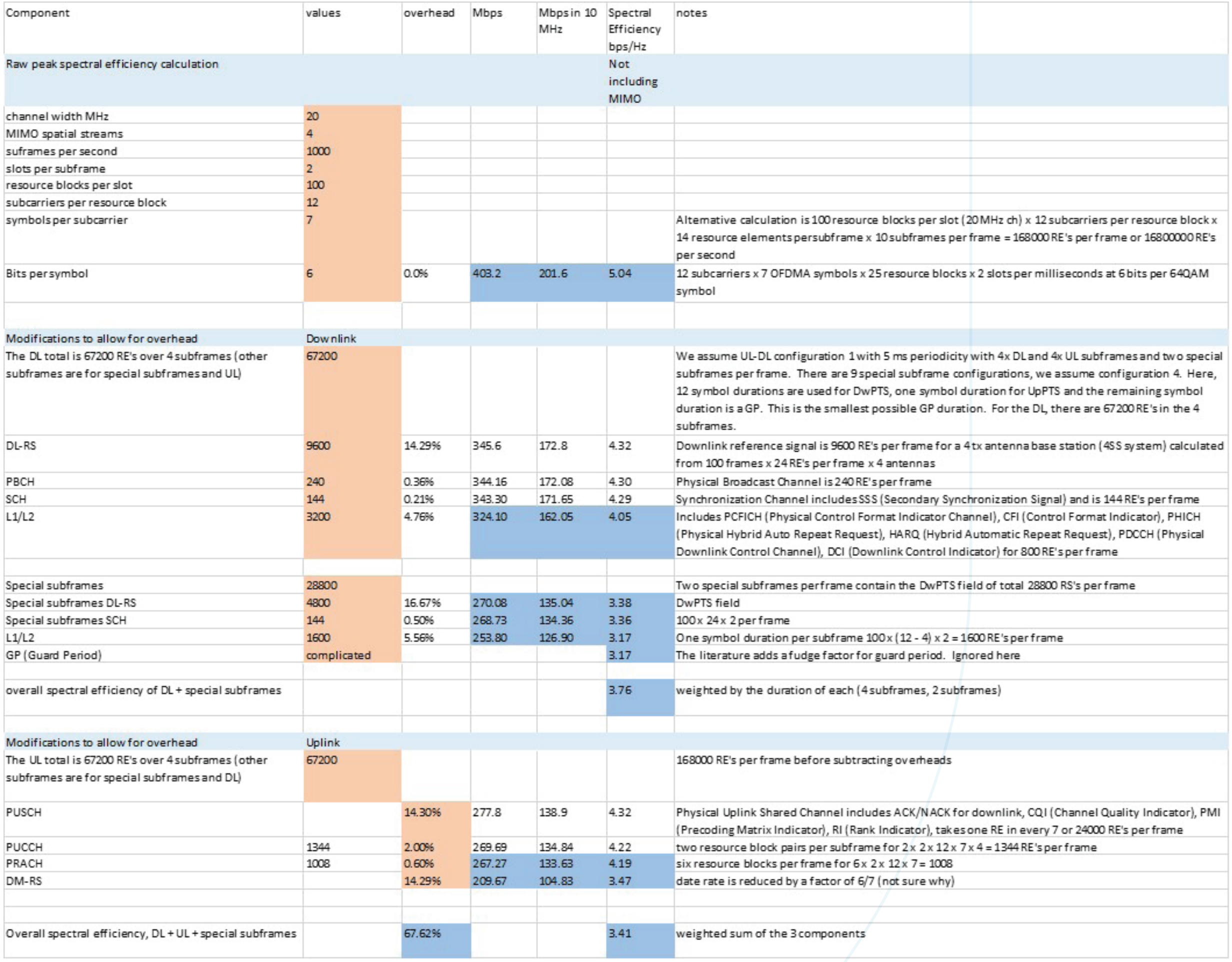
	Licensed band	5GHz band	Notes
Duplex (LAA)	FDD	FDD	Assumes for downlink only
Duplex (MulteFire)	---	TDD	Downlink + uplink
Control channel overhead (LAA)	yes	no	Control traffic is in the 'anchoring' licensed band
Control channel overhead (MulteFire)	---	yes	All traffic is in 5GHz
Listen before talk	no	yes	LBT is the sharing protocol for the unlicensed band
Co-channel interference	varies	yes	This is interference from overlapping networks using the same protocol. In LTE networks with low frequency-reuse factors, co-channel interference can be significant
Interference	not much	yes	This is dis-similar interference or noise, from other unlicensed transmitters

LTE in licensed bands typically uses FDD (frequency-division duplex) format. There are separate RF channels for the uplink and downlink, separated in spectrum so that both ends can transmit and receive simultaneously. When LTE moves into 5 GHz as LAA, it will bond the downlink only with the 5 GHz channel, so transmission is in a single direction (although there will be future versions of LAA that allow bidirectional transmission in the 5 GHz band, we only deal with downlink transmission in this paper). The anchoring licensed band connection also contains all the control traffic required for the transmission, so this overhead is not added to the 5 GHz stream.

MulteFire, with no anchoring licensed connection, must allow for both uplink and downlink traffic in a single channel at 5 GHz, so it must switch to time-division duplex operation, introducing some overhead for coordination. The control traffic is also carried in 5 GHz, adding additional overhead.

Another aspect of the 5 GHz band is the need for LBT (listen-before-talk behavior). Since the band is unlicensed and anyone is able to transmit on it, provided they meet basic underlying requirements (from the FCC in the US, ETSI and national regulators across Europe and the rest of the world) the industry has developed de-facto standards (ETSI specifies more than the FCC here) where a radio should not start transmitting if it senses another transmitter on the air. This implies that transmissions must be bounded in time, so when one station finishes, another that is waiting to transmit can proceed. These pauses and transmission intervals introduce overhead that is not necessary in licensed bands, where by definition only one operator has access to a particular channel in a particular place.

A calculation of spectral efficiency in TDD mode:



Component	values	overhead	Mbps	Mbps in 10 MHz	Spectral Efficiency bps/Hz	notes
Raw peak spectral efficiency calculation					Not including MIMO	
channel width MHz	20					
MIMO spatial streams	4					
suframes per second	1000					
slots per subframe	2					
resource blocks per slot	100					
subcarriers per resource block	12					
symbols per subcarrier	7					Altemative calculation is 100 resource blocks per slot (20 MHz ch) x 12 subcarriers per resource block x 14 resource elements persubframe x 10 subframes per frame = 168000 RE's per frame or 16800000 RE's per second
Bits per symbol	6	0.0%	403.2	201.6	5.04	12 subcarriers x 7 OFDMA symbols x 25 resource blocks x 2 slots per milliseconds at 6 bits per 64QAM symbol
Modifications to allow for overhead	Downlink					
The DL total is 67200 RE's over 4 subframes (other subframes are for special subframes and UL)	67200					We assume UL-DL configuration 1 with 5 ms periodicity with 4x DL and 4x UL subframes and two special subframes per frame. There are 9 special subframe configurations, we assume configuration 4. Here, 12 symbol durations are used for DwPTS, one symbol duration for UpPTS and the remaining symbol duration is a GP. This is the smallest possible GP duration. For the DL, there are 67200 RE's in the 4 subframes.
DL-RS	9600	14.29%	345.6	172.8	4.32	Downlink reference signal is 9600 RE's per frame for a 4 tx antenna base station (4SS system) calculated from 100 frames x 24 RE's per frame x 4 antennas
PBCH	240	0.36%	344.16	172.08	4.30	Physical Broadcast Channel is 240 RE's per frame
SCH	144	0.21%	343.30	171.65	4.29	Synchronization Channel includes SSS (Secondary Synchronization Signal) and is 144 RE's per frame
L1/L2	3200	4.76%	324.10	162.05	4.05	Includes PCFICH (Physical Control Format Indicator Channel), CFI (Control Format Indicator), PHICH (Physical Hybrid Auto Repeat Request), HARQ (Hybrid Automatic Repeat Request), PDCCH (Physical Downlink Control Channel), DCI (Downlink Control Indicator) for 800 RE's per frame
Special subframes	28800					Two special subframes perframe contain the DwPTS field of total 28800 RS's per frame
Special subframes DL-RS	4800	16.67%	270.08	135.04	3.38	DwPTS field
Special subframes SCH	144	0.50%	268.73	134.36	3.36	100x 24x 2 per frame
L1/L2	1600	5.56%	253.80	126.90	3.17	One symbol duration per subframe 100x (12 - 4) x 2 = 1600 RE's per frame
GP (Guard Period)	complicated				3.17	The literature adds a fudge factor for guard period. Ignored here
overall spectral efficiency of DL + special subframes					3.76	weighted by the duration of each (4 subframes, 2 subframes)
Modifications to allow for overhead	Uplink					
The UL total is 67200 RE's over 4 subframes (other subframes are for special subframes and DL)	67200					168000 RE's per frame before subtracting overheads
PUSCH		14.30%	277.8	138.9	4.32	Physical Uplink Shared Channel includes ACK/NACK for downlink, CQI (Channel Quality Indicator), PMI (Precoding Matrix Indicator), RI (Rank Indicator), takes one RE in every 7 or 24000 RE's per frame
PUCCH	1344	2.00%	269.69	134.84	4.22	two resource block pairs per subframe for 2x 2x 12x 7x 4 = 1344 RE's per frame
PRACH	1008	0.60%	267.27	133.63	4.19	six resource blocks per frame for 6x 2x 12x 7= 1008
DM-RS		14.29%	209.67	104.83	3.47	date rate is reduced by a factor of 6/7 (not sure why)
Overall spectral efficiency, DL + UL + special subframes		67.62%			3.41	weighted sum of the 3 components

We conclude that, considering only peak rates and error-free conditions, LTE-FDD is ~50% more spectrally efficient than Wi-Fi, and LTE-TDD is ~20% more spectrally efficient.

MODIFYING BEST-CASE SPECTRAL EFFICIENCY SCENARIOS FOR THE REAL WORLD

The figures above are based on the highest rates attainable according to the specifications, but of course when systems move into the real world there are many effects to contend with, all of them reducing performance.

Bursty and unbalanced traffic

In Wi-Fi, which was designed from the beginning for data traffic, the unit of transmission is a frame, carrying a data packet (or part of a data packet). Each frame is treated individually, so a ping-pong exchange of frames between endpoints is essentially no different from a stream from one to another. (There's actually a small amount of overhead saved by reducing inter-frame times when a stream is sent).

However, LTE – although packet-oriented – is a scheduled system and the base station must allocate Resource Blocks both per-device and (in a TDD system) for the uplink-downlink balance. Resource Block re-configuration can occur quickly but not instantly. This means that in a real-world system, where terminal devices send and receive bursty data traffic in unpredictable, time-varying patterns, there will be a certain amount of unused bandwidth resources (with conservative scheduling, or a certain amount of dropped data with over-optimistic scheduling), reducing the overall spectral efficiency.

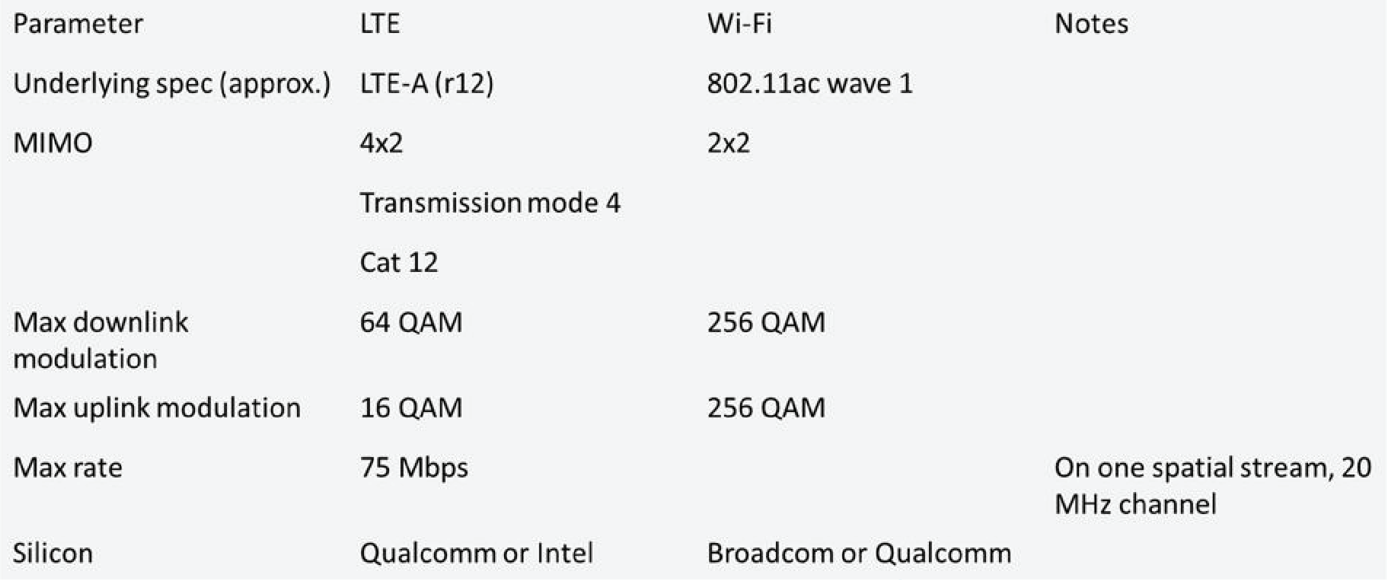
Quality of Service

The mechanisms for implementing QoS differ between LTE and Wi-Fi, but on both, overall throughput is reduced as resources are reserved for high-priority traffic, while lower-priority traffic cannot always use all available bandwidth. In Wi-Fi this is a small effect, as the CSMA/CA contention backoff of low-priority traffic is increased; in LTE the effect is more significant, as more resources must be reserved.

Practical Implementations

The calculations above are based on snapshots of the standards, comparing the best-case capabilities on paper at a point in time. We can check the results by looking at how standards are embodied in a state-of-the art device, the iPhone 7 Plus.

We immediately see some irregularities in the LTE implementation. Although the phone was shipped in 2016, it is based on 3GPP LTE r12. But it's not built to the maximum available performance as specified in r12, it is a 'category 12' device, but it falls short of category 12 specs in both uplink and downlink modulation rates. And the detail is even more complex, as in this case the Qualcomm chip used in some iPhones is more capable than Intel, but Apple restricted the Qualcomm rates to be comparable... except for the Telstra model (Qualcomm only), which can use higher rates.



Parameter	LTE	Wi-Fi	Notes
Underlying spec (approx.)	LTE-A (r12)	802.11ac wave 1	
MIMO	4x2	2x2	
	Transmission mode 4		
	Cat 12		
Max downlink modulation	64 QAM	256 QAM	
Max uplink modulation	16 QAM	256 QAM	
Max rate	75 Mbps		On one spatial stream, 20 MHz channel
Silicon	Qualcomm or Intel	Broadcom or Qualcomm	

Specifically, the QAM modulation for LTE lags the Wi-Fi capability.

The maximum data rate attainable on an iPhone 7 Plus on LTE is 150 Mbps on the downlink and only 100 Mbps for the uplink, while the comparable figures in Wi-Fi are 433 Mbps each way. When adjusted to comparable 20 MHz channels (and single-stream MIMO) the figures for LTE are 150/100 Mbps, and Wi-Fi 108/108 Mbps. This is the equivalent of the 'relative spectral efficiency' results above, before 'real world' adjustments, so it is a good match.

Metrics for comparison also spread into best- and worst-case results. These take many dimensions, but the most obvious is in maximum-average-minimum rates. While the maximum rate is often used for headlines, it is equally often unattainable given the spacing of infrastructure radios and clients. Enterprise WLANs use densely-packed access points working at low power to ensure that all clients are covered with a strong signal to give a high rate. This keeps up the worst-case rate, well above the theoretical minimum, and maintains a very high network capacity, but at the expense of considerable coordination between access point to ensure minimal interference, and advanced techniques to steer clients to the best access point.

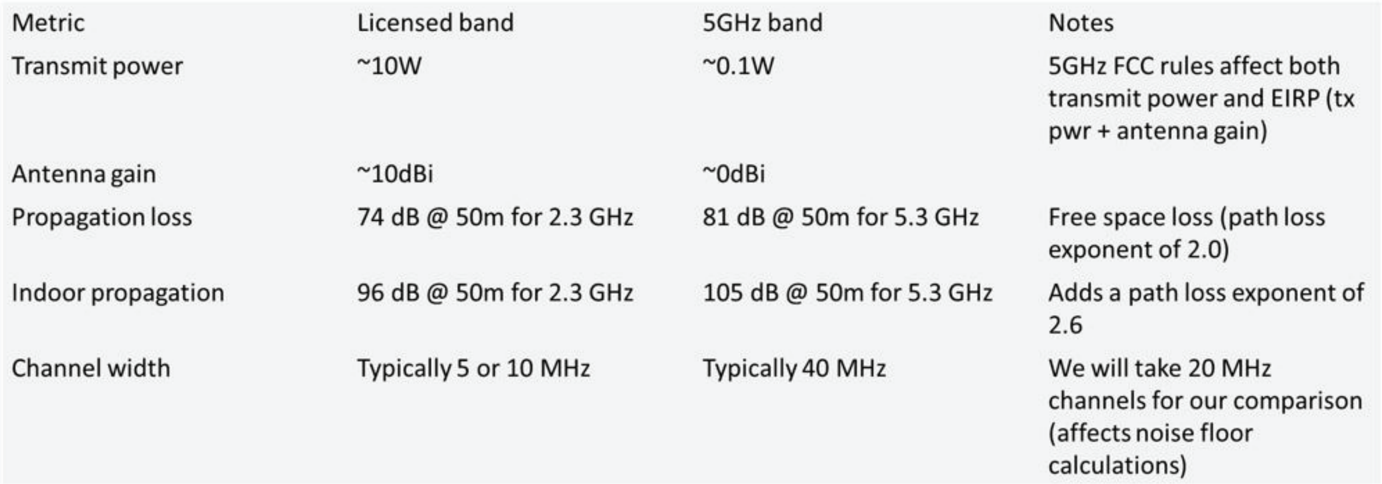
Until we get a good view of how an indoor enterprise LAA or MulteFire networks will be architected, it is nearly impossible to get meaningful figures for average rates and network capacity. And it should be obvious that maximum range is not a significant parameter in such networks.

Translation to the 5 GHz bands

Since the figures above are best-case at short ranges and without interference, they are not dependent on many factors that are band-specific. We are able to take LTE and Wi-Fi performance without considering that today, LTE is deployed in lower-frequency spectrum than the 5 GHz unlicensed band that Wi-Fi already occupies, and LAA and MulteFire will use in the future.

But moving to 5 GHz (and indoors, as our focus here is on enterprise communications that are currently served by WLANs) will entail several changes that significantly impact practical rate-and-range performance.

Some of the parameter values that change as we move LTE to the 5 GHz band:



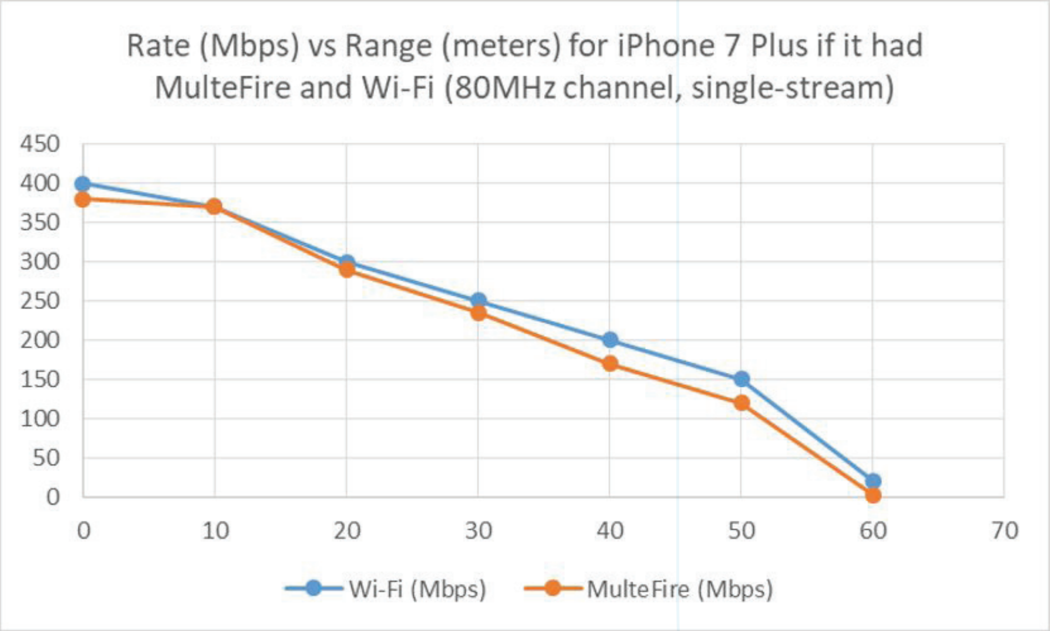
Metric	Licensed band	5GHz band	Notes
Transmit power	~10W	~0.1W	5GHz FCC rules affect both transmit power and EIRP (tx pwr + antenna gain)
Antenna gain	~10dBi	~0dBi	
Propagation loss	74 dB @ 50m for 2.3 GHz	81 dB @ 50m for 5.3 GHz	Free space loss (path loss exponent of 2.0)
Indoor propagation	96 dB @ 50m for 2.3 GHz	105 dB @ 50m for 5.3 GHz	Adds a path loss exponent of 2.6
Channel width	Typically 5 or 10 MHz	Typically 40 MHz	We will take 20 MHz channels for our comparison (affects noise floor calculations)

It is clear that the range of LTE signals in a macro cell, in the order of 1-3 km, needs drastic adjustment if we are to predict the range of an indoor LAA or MulteFire small cell working in the 5 GHz band. If we estimate 20 dB transmit power reduction and 10 dB for antenna gain, and increase propagation losses by 10 dB for an indoor office environment at 5 GHz, we can predict a reduction in range of 32x to 100x, so the 1-3 km is adjusted down to somewhere between 25-100 m.

A rigorous calculation would, among other items, treat uplink and downlink separately, as they are more unequal in licensed LTE than for 5 GHz.

All of this serves to explain why predictions of the performance of LTE protocols in the 5 GHz band depend on many assumptions. There is extensive knowledge of these effects in the academic and research community, and an expert would build a much more detailed and accurate model than the one above, but that is perhaps the point – claims for the performance of LTE in the 5 GHz band are based today on mathematical models and small-scale tests, the assumptions and conditions behind these critically affect the results, and any result must be treated as an approximation. Under these conditions, a claimed advantage of 2x range should be treated with skepticism until commercial equipment can be tested under many different conditions.

A view of expected performance (MulteFire and Wi-Fi) as it would be seen in the 5 GHz band:



A NOTE ON INTERFERENCE

Interference is a constant concern in wireless networking. This section will explore the various sources of interference that can affect Industrial IoT networks and show how these settings influenced the evolution of cellular and Wi-Fi standards.

Public cellular networks

First, consider the original setting for 3GPP specifications: public cellular networks. Here, an operator has a license to use a given channel or set of channels in a particular location. No other 'intentional' transmitters should be present on that channel. But the operator has to light up multiple cells, so in 2G/3G days narrow-beam sectorized antennas and different channels were used to avoid inter-cell interference as much as possible.

As 4G rolled out, the data requirements drove operators to use wider channels for higher data-rates, so the norm is now to design a 're-use factor' of 1, where every cell uses the same channel. This means there can be considerable inter-cell interference at cell edges, which causes data rates to degrade.

But the only forms of interference expected in a network using a licensed channel are random noise events from electrical machinery and other such sources, and inter-cell interference that is from other transmitters controlled by the operator. The latter is 3GPP-on-3GPP interference and can be managed within the operator's organization.

Private 4G/5G networks will have similar interference profiles to public cellular networks, although the issues may be simpler due to a smaller number of cells and transmitter sites.

The limited interference environment is one reason the 3GPP chose a centrally-coordinated, time-division-multiplexing scheme for its transmissions, where the AP transmits a time reference that defines timeslots, and downlink traffic to different clients is assigned to different timeslots.

Most cellular deployments use FDD (Frequency Division Duplex) where the uplink and downlink use different frequency channels, avoiding coordination between uplink and downlink. Newer standards, including 4G/5G, have a TDD (Time Division Duplex) option where uplink and downlink share the same channel. This introduces some complexity, because as the balance of uplink-to-downlink traffic changes, timeslots must be re-assigned from one function to the other if spectral efficiency is to be maintained.

Cellular networks in general expect a benign interference environment.

Wi-Fi 5 networks

Wi-Fi standards borrowed many of their original techniques from wired Ethernet, in particular the CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) protocol makes packet-by-packet transmission decisions across a number of un-coordinated clients, where each client with a packet to send first senses whether the air is clear, transmitting if it is, or if it sees another packet on the air, waiting a random time interval before trying again. Every packet gets an immediate ack response from the receiver.

The packet-by-packet, uncoordinated protocol is well-suited to bursty data, where the volume of traffic each node transmits can vary instant by instant. It does not matter whether a sequence of packets is generated by one node or across many, the protocol handles these cases in identical manner. TDM protocols, in contrast, are not able to quickly react to changes in traffic patterns, either for single clients or across populations, as the allocation of timeslots must be changed to increase or decrease the capacity available to different clients. One solution is to leave headroom by allocating more bandwidth than needed, but that contributes to spectral inefficiency.

This Wi-Fi 5 protocol is very robust to errors – if there's no ack, the packet is retransmitted – and because there is no central control or coordination function, it can easily handle overlapping cells, where several access points choose to transmit on the same channel in the same area. Naturally, performance declines in these situations, as can be experienced in urban areas or sometimes in multi-family apartment blocks, where many Wi-Fi access points are crammed in a small space. But even though performance can suffer somewhat, Wi-Fi networks continue to function under such conditions with no overarching management or inter-network coordination.

This robustness to interference, both general electromagnetic noise and other Wi-Fi signals, is a key attribute of Wi-Fi, well-matched to its un-planned deployment model. But it causes some difficulties in the domains of QoS and network capacity.

For QoS, a loosely-coupled, uncoordinated system lacks determinism: there is no control signal to indicate that because one node has higher-priority traffic to send, others should allow it to transmit. Wi-Fi solved this problem by adjusting the timers used in random backoff delays, so higher priority traffic can use shorter timers and transmit ahead of other traffic. This scheme, known as WMM (Wireless MultiMedia) works well at low traffic loads but can break down when activity levels exceed 90% of available airtime.

CSMA/CA also has considerable overhead, as packets are not coordinated to follow each other immediately. There is a time-gap for the contention window while backoff timers run out and nodes can sense the air is free and start a new transmission.

Comparing cellular and Wi-Fi susceptibility to interference

As touched on above, the cellular protocols started with a relatively inflexible TDM structure, which was appropriate for voice traffic of predictable bit-rates in the absence of interference.

As data traffic became more important, it had to make provision for the burstiness and re-transmissions required, and did so by attempting to speed up the mechanisms for re-allocating bandwidth per-client, adding several levels of forward error correction and retransmission of which HARQ (Hybrid Automatic Repeat Request) is the latest. We noted above the requirement to re-balance uplink and downlink capacity: this is possible in TDD mode, but not FDD where each direction is bounded by a separate, paired RF channel.

Meanwhile, Wi-Fi started with a very robust but relatively inefficient (of airtime or spectral efficiency) MAC structure, CSMA/CA, and proceeded to modify it for QoS as noted. It also allowed nodes with multiple packets to send them in a stream and receive directed acks, in the Block-Ack protocol, increasing spectral efficiency, especially for high-bandwidth streams such as streaming video.

This abridged history shows that, starting from different initial positions but subject to the same pressures, the two standards have been converging for some time. The latest convergence is the adoption, in Wi-Fi 6, of OFDMA, a multi-user protocol where the access point can coordinate all client transmissions through the use of trigger frames, and direct transmissions to sub-channels in the frequency domain.

The increased levels of control are very significant. If it wishes, a Wi-Fi 6 access point can schedule transmissions into the future for both uplink and downlink traffic and assign each client an arbitrary level of performance by specifying data-rate and other parameters. This allows control of QoS, including latency and jitter, also error rate behavior in the event that an access point is overloaded and must shed traffic. In this scenario, Wi-Fi 6 is in practice equivalent to 4G/5G TDD.

With Wi-Fi 6 now able to offer a range of behavior from the earlier packet-by-packet transmit decisions to full OFDMA with central control in different combinations, the opportunities to tailor network behavior to wireless conditions are much broader. Full OFDMA is most suited to carrying relatively stable network traffic with the best efficiency and most deterministic QoS in a benign RF environment, while packet-by-packet decisions may continue to prove more resilient in situations with overlapping, uncoordinated access points and higher levels of interference.

Interference may appear from different sources. First, there is co-channel interference, where overlapping networks of the same type are set up on the same RF channel. This is common in public areas of cities or apartment blocks that may be in range of dozens of Wi-Fi networks. Wi-Fi has, from the beginning, used the CSMA/CA protocol to limit transmit times and ensure that stations waiting to transmit all have a fair chance of getting on the air, whether part of the same network or an overlapping one on the same channel. CSMA/CA adds considerable overhead to Wi-Fi but enables its great resilience, ensuring that, even though performance for each user is degraded when networks share spectrum, they all share the air with fairness and use the available channel capacity for the best overall performance.

It is not yet clear in practice how well the LAA or MulteFire protocols will deal with overlapping networks (whether using Wi-Fi or other protocols) sharing the same channel but set up by different operators. Most of the work to date has been to investigate how LAA and MulteFire interact with Wi-Fi transmissions. Overlapping communications networks at 5 GHz such as cordless phones, baby monitors and point-to-point links are another class of interference that includes not just LAA, MulteFire and Wi-Fi but also other legal users of the 5 GHz band.

Added to this, interference is generated by unintentional transmitters of RF noise such as sparking machinery, vehicles, lighting units, microwave ovens and other sources. This interference is typically shorter in duration and more intermittent.

All sources of interference degrade transmission by causing errors in detection. In 4G/5G and Wi-Fi, this results in missed frames and retransmissions, and often in a subsequent reduction in data rates to avoid further frame loss, which lowers throughput and reduces spectral efficiency.

Dealing with interference requires a comprehensive listen-before-talk protocol; variation in coding levels; ack, block-ack and retransmission protocols; and a rate-control algorithm. Operating in an unlicensed band also means that interference is far more common than in licensed bands, and this will stress forward-error-correction and retransmission protocols.

For the purposes of this paper, the increased levels of retransmission will mean lower effective data-rates and lower throughput. It is very difficult to predict just how much lower, but Wi-Fi's frame-by-frame transmission is more flexible and will accommodate interference better than LTE's scheduled scheme.

QUALITY OF SERVICE

Quality of Service is a broad concept covering latency, jitter, error rate, and probability of dropping traffic under overload. In this section we will characterize the significant parameters affecting QoS in 4G/5G and Wi-Fi networks for Industrial IoT.

QoS requirements are set by the traffic characteristics and the uses to which an application is put. There are several orthogonal tracks, but we will group them here with some typical application requirements.

1) Low-bandwidth IoT sensor traffic for monitoring purposes, for example temperature or vibration monitoring of a machine. This traffic is low-bandwidth, perhaps 10 kbps either steady-state or, more usually, periodic at intervals from seconds to hours. It is generally not critical that every update is received: one or two measurement failures can be tolerated; and latency and jitter do not need to be bounded in practical terms.

2) High-bandwidth signals, for example video. Where images from a manufacturing cell need to be transmitted to a data center or edge compute center for processing, there are requirements for error rate (e.g., <5% packet loss) and latency (e.g.,, <500 msec), while the bandwidth may be high (e.g., 10 Mbps).

3) Where the video example above is extended from offline analysis to real-time control, for example where a process must be modified or stopped if certain conditions are detected, or where a machine is controlled in a tight real-time loop, QoS requirements may be more stringent (e.g., 2% packet loss with 300 msec latency and 50 msec jitter for the round-trip control loop). This example illustrates that QoS is an end-to-end phenomenon: it extends across the wireless links, but also to the wired network and processing delays in the edge compute center.

4) Other traffic may be safety-critical. Here, the packet loss ratio must be very low, but the overall probability of outage must also be bounded. Examples include real-time control of moving equipment such as cranes in container ports, or monitoring of chemical plant to prevent dangerous situations. Latency and jitter may not be significant here, but typically these applications will require some level of redundancy in equipment and communications links, and assurance that this traffic will be delivered in the event of overload and other unusual scenarios.

QoS can be affected by several factors, including:

1) Processing delays or transmission delays in various network nodes: wired, wireless and computing

2) Delay and overload due to other traffic on the network

3) Interference on the radio link, from other intentional transmitters or from electrical noise

QoS in 4G/5G public networks

Where an Industrial IoT network uses the public cellular network, it benefits from the lack of interference in sheltered, licensed spectrum. This is a significant benefit, but does not provide complete protection against electromagnetic interference from unintentional transmitters.

Apart from electrical noise, the public network is generally considered reliable for data traffic. However, in the event of overload there is – in today's cellular networks – no guarantee that certain connections will be given priority while others will be dropped. This may be remedied as mobile operators move to 5G architectures and implement network slicing, which will reserve network resources for preferred customers and applications.

Similarly, there is no guarantee of latency and jitter over the public cellular network today, but it is generally good: within 40-80 msec for most connections. With improved radio techniques in 5G, this will be reduced to a design goal of 10 msec, at least for the wireless link. At this point, wired network and compute latency will become significant, so the overall latency-jitter envelope may still be closer to 50 msec than 10 msec.

QoS in private 4G/5G networks: licensed and unlicensed spectrum

Private 4G/5G networks offer more opportunity for the customer to control the QoS per-application, provided the equipment designer offers APIs to configure this, as the customer is responsible for managing the network.

4G and 5G use framed time-division and frequency-division multiplexing with central control from the base station that allocates RBs (Resource Blocks) to each client device for both downlink and uplink. The framed, controlled structure ensures that traffic from one stream does not impinge on others.

In addition, 4G/5G assigns a traffic stream to a QoS class including GBR (Guaranteed Bit Rate), non-GBR (non-Guaranteed Bit Rate), Priority Handling, Packet Delay Budget and Packet Error Loss rate, through a mechanism called QCI (QoS Class Identifier). The 4G/5G standards list 17 standardized QCI classes, including 'conversational voice', 'mission critical video user plane', 'low latency eMBB applications'. QCI extends into the packet core for end-to-end QoS. Configuration of these QCI levels is unlikely to be available in 4G equipment for private networks, but may emerge with 5G equipment.

Private 4G/5G networks in licensed spectrum will benefit from guarantees of non-interference, but in unlicensed bands there will be no such guarantee.

QoS in Wi-Fi 5 and Wi-Fi 6 networks

The different mechanisms used for QoS in the Wi-Fi standards were described briefly in the section on interference, above. Up to Wi-Fi 5, the CSMA/CA channel access protocol was modified to allow nodes with high-priority traffic to gain priority access over others. This WMM protocol defines four classes: 'Voice', 'Video', 'Best Effort' and 'Background'. These are mapped to the DSCP (Differentiated Services Code Point) and 802.11p priorities used in LAN equipment and routers, for end-to-end priority.

WMM is effective in prioritizing traffic at up to 90% network load, but can break down in overload conditions, as it is not always possible to suppress low-priority nodes on the uplink. However, since access points handle all downlink traffic, they are able to drop low-priority traffic and maintain the high-priority streams, and where acknowledged protocols like TCP/IP are used, this is an effective method of managing priorities. Wi-Fi 5 is generally accepted to have jitter-latency in the 50 – 150 msec range, reduced to 30 – 70 msec for high-priority WMM classes.

With Wi-Fi 6, the new OFDMA capability allows comprehensive control of traffic from the access point, in both downlink and uplink directions. Not only are transmit opportunities controlled, but OFDMA also allows traffic to be assigned to dedicated RUs (Resource Units, equivalent to 4G/5G Resource Blocks) for deterministic QoS characteristics. OFDMA will allow jitter-latency to be reduced to the sub-10 msec level: it can improve considerably on Wi-Fi 5 by offering more frequent transmit opportunities than the traditional packet-by-packet contention.

Since Wi-Fi uses unlicensed spectrum, it may encounter interference from other users of the band, as well as general electromagnetic interference. This can be mitigated in industrial areas that are remote from the public, as private companies can institute policy for use of Wi-Fi on-campus, and WLAN equipment is able to monitor 'rogue' access points that are not part of the company network, but the possibility of interference is higher than for licensed spectrum.

Conclusion: Quality of Service

While many Industrial IoT network requirements may be satisfied by existing Wi-Fi QoS classes, there will be cases where at least some of the parameters implicit in QoS must be bounded. In the short-term, this will require custom configuration of private 4G or Wi-Fi 5 network equipment, as well as the LAN and computing infrastructure. Edge computing will play a significant part in reducing latency and jitter and the probability of outages, as will TSN (Time Sensitive Networking), a series of new Ethernet LAN/MAN standards from the IEEE 802.1 working group. In the medium-term, Wi-Fi 6 will offer significant new opportunities for the Industrial IoT network to implement 'hard' QoS, and 5G public and private networks will deliver equivalent functionality.

RANGE CONSIDERATIONS

Range is important in wireless systems because it determines how many base stations are required to cover a given area, which has implications for network cost. Wireless signals carry for great distances – the primary challenge when building a high-capacity WLAN is from inter-cell interference that bleeds into adjacent cells – therefore, to properly estimate range it is necessary to specify a performance objective at the cell edge. This is usually expressed as the minimum data-rate acceptable for downlink/uplink to IoT devices. Modern wireless systems reduce the data-rate as signal-to-noise ratio falls, to maintain a reasonable error-rate.

Other parameters, such as probabilistic estimates of availability and fading interference outages can be significant, especially for outdoor networks and ultra-reliable networks. In a network with high-rate traffic, the capacity of an individual infrastructure radio may be the limiting factor in network design, but in most single-purpose Industrial IoT networks, effective range will be the significant measure.

Industrial IoT networks are often low-rate, and a simple analysis might seek to minimize the number of access points, but factory buildings may enclose large metal or concrete structures, and other RF obstructions. In this type of environment, larger numbers of access points operating over short ranges with small cell sizes can be superior: the environment limits inter-cell interference which boosts overall network capacity.

Conversely, in outdoor networks with sparse client populations, range-at-minimum-rate will usually be the limiting factor.

Operating frequency, allowable transmit power and antenna size and gain all limit the range of a wireless signal. Most of the differences between 4G/5G and Wi-Fi range stem from these parameters, as licensed frequencies are usually low-band, high transmit power is allowed by regulators, and external tower-mounted antennas can be large. In contrast, Wi-Fi's transmit power is limited by regulation, the 5 GHz unlicensed band is higher-frequency than most 4G bands, and wall-mount, indoor access points with internal antennas must be a handy size. Outdoor point-to-point access points are usually fitted to external, high-gain antennas.

Range of 4G/5G connections

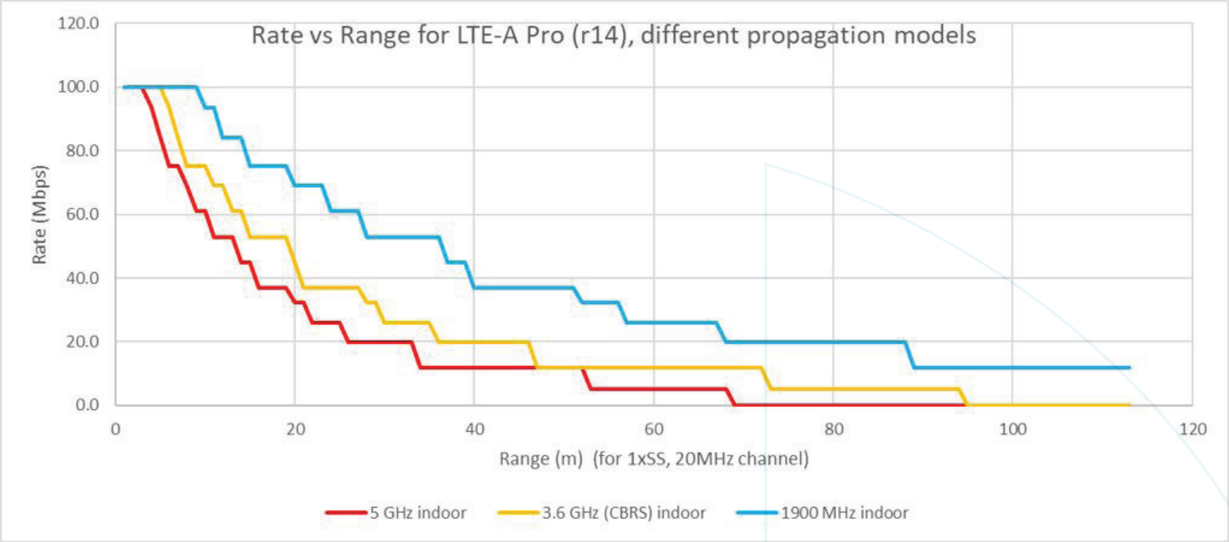
Most 4G/5G base stations are deployed outdoors, for the macro network. They support cells of the order of 1 – 3 km radius from the tower. But these are full-power transmitters with large, sector antennas and are not suitable for indoor deployments. We can make the following coarse characterization for 4G/5G radios:

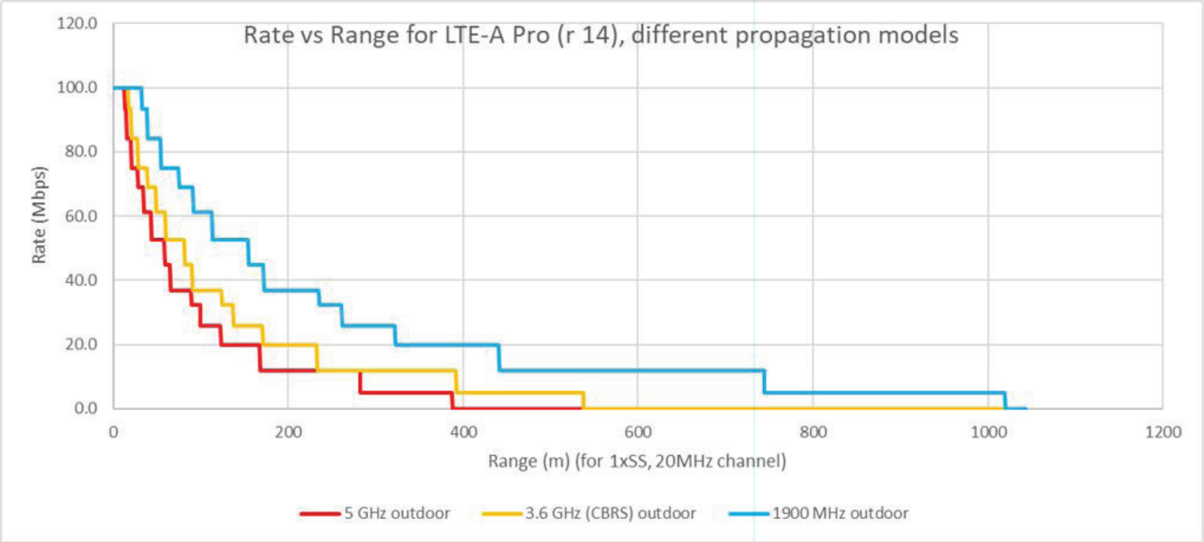
1) Indoor femtocell, 4G; 17 dBm tx power with omni antenna; 700 MHz operating frequency; 16 clients @150/50 Mbps max dn/uplink... Indoor range ~50 m; AP spacing 50 meters for 75/25 Mbps cell edge rates

2) Indoor femtocell, 4G; 17 dBm tx power with omni antenna; 2600 MHz operating frequency; 16 clients @150/50 Mbps max dn/uplink... Indoor range ~25 m; AP spacing 30 meters for 50/25 Mbps cell edge rates

3) Outdoor microcell, 4G; 37 dBm tx power with omni antenna; 2300 MHz operating frequency; 32 clients @ 110/10 Mbps max dn/uplink... Outdoor range ~1000 m; AP spacing 1500 m for 50/25 MHz cell edge rates

Note that, due to the asymmetrical links when the base station has much higher transmit power than the client, 4G/5G downlink rates are ~2x uplink rates [this may become accentuated for IoT sensor devices, where power, packaging and antenna size issues may force design compromises].





Range of Wi-Fi connections

Wi-Fi access points operate with a regulatory maximum ~23 dBm indoors and ~30 dBm outdoors (specified as EIRP, so higher-gain antennas force reduced transmit power).

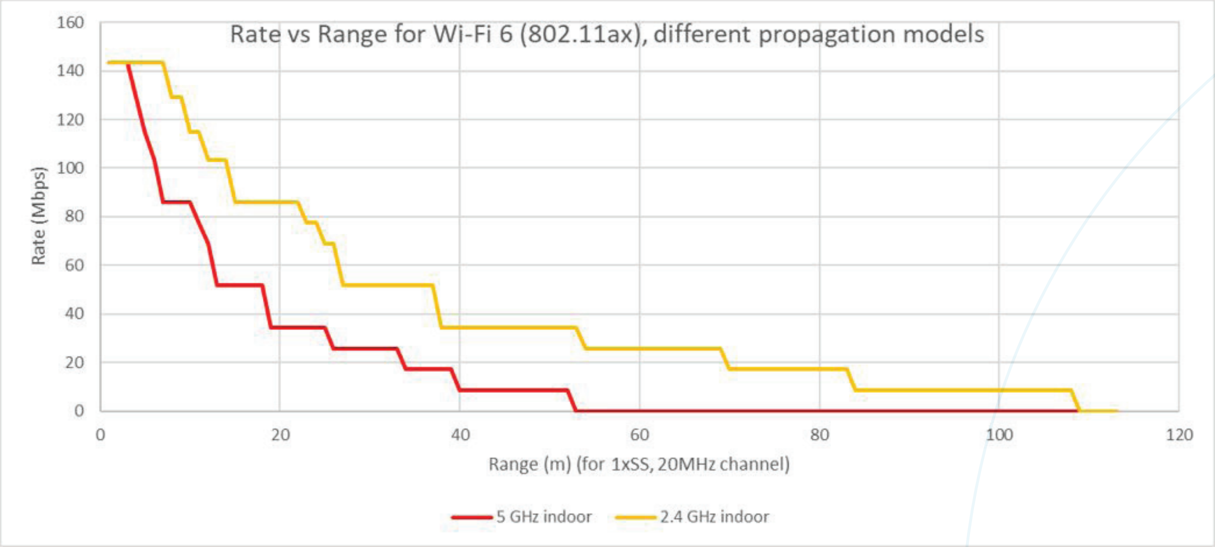
In the 5 GHz band, the wavelength is 6 cm so antenna size can be small, and modern Wi-Fi 5 access points include 4 antennas for each of two radios, for a total of 8 antennas (Wi-Fi 6 access points will scale to 12 antennas). Indoor access points are built for area coverage with omnidirectional or pancake coverage zones, and are optimized for wall- or ceiling-mounting.

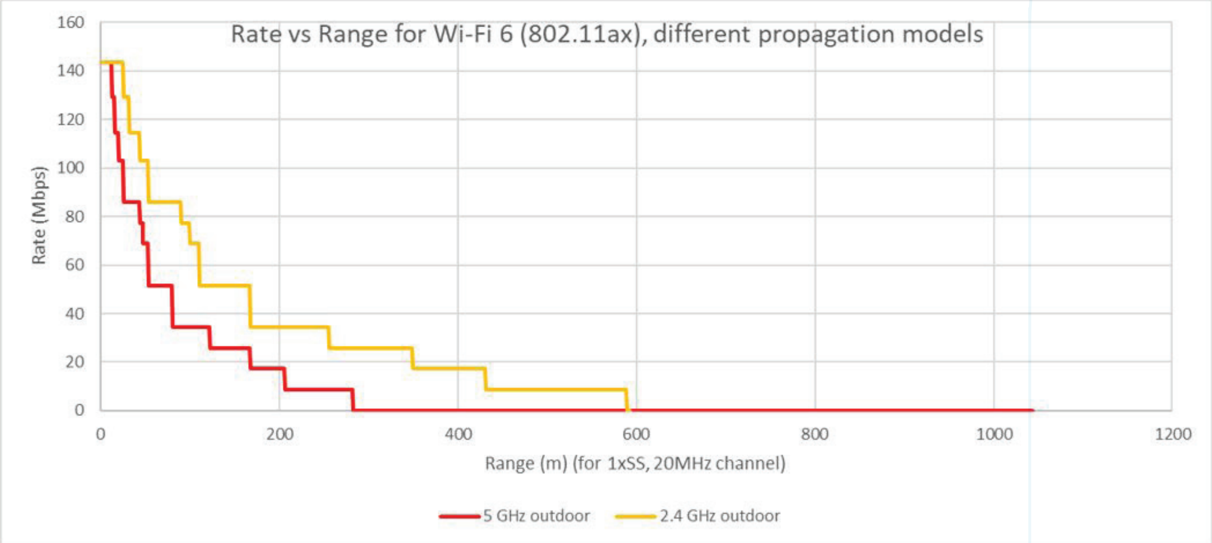
The following are general rules of thumb for Wi-Fi coverage; for factory and process plant installations, a site survey by a knowledgeable WLAN engineer is recommended.

1) Indoor access point, Wi-Fi 5; 20 dBm tx power with omni antenna; 5200 MHz operating frequency; 300 clients @500/500 Mbps max dn/uplink... range 15 - 25 m; AP spacing 15 - 25 m for 100/100 Mbps cell edge rates, 100% coverage overlap [a little more in the 2.4 GHz band].

2) Outdoor access point, Wi-Fi 5; 20 dBm tx power with omni antenna; 5200 MHz operating frequency; 300 clients @500/500 Mbps max dn/uplink... range 100 - 300 m; AP spacing 500 m for 100/100 Mbps cell edge rates, 70% coverage overlap [a little more in the 2.4 GHz band].

Outdoor point-to-point range 700 - 3000 meters for rates 200 - 800 Mbps bidirectional with high gain antennas. Wi-Fi rates are bidirectional as the client devices are usually as capable as access points. This may change for IoT sensor devices, where power, packaging and antenna size issues may force design compromises, but holds true up to Wi-Fi 5.





Conclusion: Range considerations

We are accustomed to the idea that a 4G cell is in the order of 1-5 km in diameter, but current 4G femtocell equipment is forced into many compromises as it shrinks radios from macro base-station proportions to small-cell scale. Despite this, indoor femtocells operating in low-band licensed frequencies are generally expected to have twice the range of equivalent Wi-Fi access points. However, the data rates supported by these units are lower than for Wi-Fi, especially for uplink data. Unlike Wi-Fi, 4G equipment in practical incarnations does not provide symmetrical data rates – uplink rates are typically half the headline downlink rates, but for IoT sensors it is the uplink that is critical.

For outdoor coverage, 4G equipment operating in low-band licensed frequencies has considerably farther reach than Wi-Fi in coverage situations, at the expense of equipment cost and size. However, Wi-Fi can use point-to-point links and form mesh networks (see later in this paper) to counter this advantage in some situations.

As 4G/5G networks move into mid-band and high-band frequencies, and especially as they move to the unlicensed 5 GHz band and become subject to the same regulations as Wi-Fi, range performance will converge.

Network topology building blocks for private 4G/5G and Wi-Fi industrial networks

For both technologies, the building blocks are similar. Indoor, short- and intermediate-range coverage networks in factories and similar environments will deploy a grid of infrastructure radios (base stations or access points), spaced to provide the appropriate minimum data rates at cell edge, with some level of cell overlap for redundancy and coverage assurance: when the path to one access point is interrupted, there will be a second path for the client to switch to. These grids may be modified to adapt to local conditions: metal structures, open corridors, concrete walls and others, with specialized external antennas or custom mounting arrangements, but the default selection will be wall- or ceiling-mounting of standard radio units with internal omni antennas. Backhaul from edge radios to the core network will normally be by copper or fiber Ethernet, although mesh, point-to-point wireless, or even the public cellular network are all options.

1) Where the network serves only low-rate Industrial IoT devices, the emphasis will be on providing assured coverage. Generally, the larger the number and density of access points, the more reliable the connection: in Wi-Fi, clients will seek out the strongest signal, and switch access points when conditions change. There is less history of this with indoor 4G equipment: the femtocell industry has a more limited track record of tightly-coordinated radio infrastructure than Wi-Fi, so although 4G inter-cell handover is readily supported in the standards, its performance should be tested with the particular IoT equipment selected.

2) Many of the emerging 'Industry 4.0' applications involve video and image processing where data is streamed from manufacturing cells to a data center, often at high-definition, and these networks require more careful planning to ensure that capacity is available for all required applications. But the principle is similar to other indoor networks.

3) As the focus moves from relatively bounded indoor areas to longer-range outdoor areas, up to ~500 - 1500 m across, network models become more complex. Siting of infrastructure radios is important: usually, the higher the better for good area coverage, but there is likely to be more shadowing from metal structures and buildings that require outside-in coverage, so more detailed site surveys will be necessary. Copper or fiber backhaul may be available at preferred mounting points, but wireless backhaul links, or mesh connections are often required. Wi-Fi has more experience with mesh and point-to-point hybrid backhaul, as 4G in the public network would generally install a single base station with coverage over a wider area. But this model will be challenged as frequencies move to mid-band and high-band spectrum and power restrictions limit range, so the 4G/5G network for these sites will need to develop new expertise.

4) Very wide-area networks, 3000 m across and more, may need different approaches. If the area is contiguous, and radio mounting sites are available with or without backhaul, the network design above can be expanded without limit. But if the requirement is to cover patches or islands of industrial plant surrounded by large tracts of land where there is no right-of-way to set up radio towers – perhaps a set of oil platforms spaced several km apart across public and private lands – it may make more sense to use the public cellular network, or alternatives such as satellite service for backhaul, and install local coverage at each site feeding this backhaul.

Conclusion: Network topology building blocks

In both 4G and Wi-Fi ecosystems, there are several established techniques for building coverage, capacity and backhaul networks. 4G will build on indoor femtocell experience and borrow from the public cellular network for wide-area and backhaul options with the move to 5G. Wi-Fi has extensive experience with indoor, campus-wide and large-scale outdoor networking.

Current Wi-Fi equipment supports higher data rates, particularly for the uplink, than 4G, while cellular infrastructure generally has longer range. It is not yet clear how much this will change as Wi-Fi 5 moves to Wi-Fi 6, and 4G to 5G. All these techniques will be improved as experience is gained from the emerging Industrial IoT market.

DEVICE AVAILABILITY AND ROAMING IN AND OUT OF THE PUBLIC CELLULAR NETWORK

This section investigates device availability, authentication and roaming questions that may arise as a company builds an Industrial IoT network. For networks that are used only for Industrial IoT purposes, broad device selection and roaming may not be a requirement as IoT devices will be connected only to that network.

But some companies building a private 4G/5G network will expect it to provide cellphone service for employees and visitors in addition to IoT functions, and it is important to realize that few of the models described earlier in this paper support the level of universal connectivity and roaming we are familiar with in the public cellular network.

We have many expectations of the public cellular network in terms of universal compatibility and roaming. By compatibility we mean the ability to purchase a cellular device and have it automatically connect to an operator's network. This requires devices and networks to support uniform protocols and licensed RF bands.

The ability to easily roam from one operator's network to another is another great attribute of the global public cellular system. This function is enabled by common protocols, a known set of licensed RF bands and links between networks that allow a subscriber's home operator to be contacted for authentication and policy status when roaming to a visited operator's network.

The benefits of roaming can be further broken down: first, there is access to the network which might be important to consume local resources. Second is data service to the Internet, which is not usually a problem. Third is the ability to make phone calls and SMS messages to other subscribers of the public network.

Network models: extensions of the public cellular network

When a network is built by a mobile operator, as an extension of its cellular network, nearly all of these functions should be available as one would expect.

1) Any cellular device should be able to connect, provided:

a) It is not SIM-locked to another carrier

b) It supports the licensed frequency bands used, which is expected in this case, as they will be part of the operator's standard licensed frequencies.

2) Roaming-in by subscribers of other operators will be technically supported:

a) But, depending on roaming agreements and operators' policy, the subscriber may have to manually force the device onto the network.

b) This is particularly true for national roaming, and where the subscriber's home operator provides good coverage of the area.

3) Roaming-out by subscriber devices of the network's home operator should be supported, provided they are set up as consumer subscriptions.

4) International roaming-in and -out should be well-supported.

The difficulty with this model is often the reverse of its universal service characteristic. The enterprise might prefer that while employees and company-owned devices are allowed on the network, visitors and third parties are excluded, especially where Industrial IoT is deployed. This can be difficult for the operator to accomplish, as it would require new control functions not used in the public network.

Network models: private 4G/5G networks in licensed spectrum

Where the network is 'private', roaming can become restricted on a number of levels.

First, consider device availability.

1) Although we consider the cellular system universal, it has many options, one of which is the very broad range of frequency bands allocated nationally and world-wide. In the US alone, there are 3 bands for 2G, 3 for 3G and 9 for 4G, and world-wide there are many more. Most phone makers build different variants for global markets like Asia, Japan, Europe and the US (for example, Apple has 6 regional variants of the iPhone Xs, Samsung has 8 variants of the Galaxy S9). Some phones are built for specific operators and have further restrictions.

2) This should not be an issue for mainstream devices in the US, but where new bands such as CBRS at 3.5 GHz are chosen, device availability will be limited.

3) Technology is also a risk. While most phones are backwards-compatible and today include 4G, 3G and 2G radios, small cell equipment used for Industrial IoT networks may only support one – usually 4G – level of technology.

But roaming-in may be the more significant obstacle. If the Industrial IoT network is not part of a major operator's network, there are broadly 4 models for providing 'neutral host' roaming support. ['Neutral host' usually means multi-operator roaming-in support, but the models are equally applicable here.]

1) One 'sponsor' operator that offers roaming service to others. To enable this, the company will need to make an arrangement with a major mobile operator that allows SIM-authentication of that operator's subscribers from this private network. Th ere are very few of these agreements in place today. After that, the operator must arrange for other operators' subscribers to roam-in when they come into coverage.

2) The private network may be owned and installed by a small, third-party operator which often also arranges spectrum. In this case the network will appear as "XYZ Wireless" on a cellphone. Dedicated company devices will need a custom-programmed SIM card for this network, while roaming-in subscribers of mainstream operators will rely on roaming agreements between the small operator and the majors. Very few of these arrangements are in place today, either.

3) The standards support various levels of infrastructure sharing between operators, where a single set of radio nodes, for instance, transmits multiple operators' signals as the home operator. While these networks are made possible by 4G/5G standards, operators have shown no desire to build the business relationships and processes to enable them: they exist only in theory at this point.

4) Install small cells from multiple operators, with overlapping coverage. Subscribers will see a signal from each of their home operators. This is the default configuration for DAS systems (although they will try to aggregate all RF signals onto one antenna system) to provide multi-operator support.

It must be reiterated that, although superior options such as (1) and (2) are technically feasible, major operators have to date been reluctant to support them, and very few 'real-world' examples exist.

Roaming-out for this model may be more problematic than roaming-in. Building a private 4G/5G network implies installing a dedicated, private authentication server and provisioning device credentials (usually SIM cards) against that server. But those SIM credentials will not be recognized by any other operator unless a roaming agreement is put in place (see above).

Perhaps the best option, if roaming-out is required, would be (2) above, where the network construction, operation and management is effectively outsourced to a third-party, small operator that is better positioned to negotiate the roaming agreements necessary. This is a very early-stage market in the US as of early 2019.

Network models: private 4G/5G networks in unlicensed spectrum

This final model is rather different, and more speculative than those above as no equipment or 'real-world' networks exist, as of early 2019.

The first question surrounds device availability. If these networks are built, devices are likely to be existing cellular IoT devices, modified for new frequency bands. Therefore, they will predominantly use SIM authentication. Depending on the authentication model, they will be subject to the same authentication and roaming issues as the private 4G/5G networks explored above.

It seems improbable that major operators will endorse and build private networks entirely in unlicensed spectrum, so the private network model seems most likely.

Devices may be designed explicitly for these unlicensed frequencies, in which case they will be unable to roam to any existing cellular networks, or modified from existing designs to add the new unlicensed bands.

Companies considering this model for an Industrial IoT network should assume it will constitute an island of connectivity separate from the public cellular networks.

Wi-Fi mechanisms for inter-carrier roaming

Wi-Fi has a very good record of device and technology compatibility. The latest devices purchased today will operate on access points from 15+ years ago, and vice versa. And there is global harmonization of frequencies; a traveling businessperson's Wi-Fi device will be able to connect to an access point anywhere around the world.

Authentication is also very flexible. As explained earlier in this paper, a Wi-Fi device can be authenticated using passwords, X.509 certificates or SIM cards, and when using WPA2-enterprise, the authentication server can be locally-managed or service provider-administered. While Wi-Fi services can be configured for exclusive use of certain devices, which might be appropriate for an Industrial IoT installation, they can simultaneously support other forms of authentication, such as guest access, if desired.

For visitors to an enterprise campus, manual selection and authentication or guest-registration can create friction. A recent Wi-Fi Alliance certification, 'Passpoint', allows service providers to program devices to automatically recognize and connect to third-party networks which support a RADIUS authentication path to their core networks. In this way, Passpoint enables seamless SIM-based authentication over Wi-Fi. Passpoint adoption is growing, notably in the US.

Wi-Fi devices are data-oriented, using the TCP-IP protocol for universal data and Internet connections. But making voice and text calls to public network subscribers is more difficult. The recent adoption of 'Wi-Fi Calling', a protocol that allows voice and text service over Wi-Fi, allows Wi-Fi devices to participate in roaming across the public cellular and telephone networks wherever they can get an Internet connection. Wi-Fi Calling is supported on all major mobile operating systems, and by 100+ mobile operators world-wide, as of early 2019.

Conclusion: Device selection and roaming

The range of devices available for operation on an Industrial IoT network will be a significant issue for the next few years, as the market ramps up. Regardless of the wireless standard chosen, it is important to research device capabilities in the planning stage.

While any existing or new Wi-Fi device will be capable of connecting to an industrial Wi-Fi network, the relative lack of maturity of the 4G/5G IoT market results in more limited device availability. Further, if a private 4G/5G network is planned, using a new frequency band, the range of devices will be even smaller. These issues can be minimized by using aggregation devices to avoid touching every sensor, but at the expense of complicating network design.

Roaming with the public network, in its many permutations, may or may not be an issue for many Industrial IoT customers. But it should be considered, as superficial assumptions drawn from the public cellular network may not be accurate.

MOBILITY: CLIENTS MOVING AT SPEED

Moving clients – cars and trains among others – can be a significant component of Industrial IoT networks. Networks along roads and railway lines present challenges for all wireless technologies:

1) Coverage. Building a base-station network that covers the large distances involved can be difficult, and the shorter the range of the signal, the more base stations are needed.

2) Handovers. Clients need to move from base station to base station. This always involves some disruption, and is exacerbated by short-range links: as inter-base station distances shrink, handovers become more frequent.

3) Radio limits at speed. While this is perhaps the least significant of the challenges, it can be limiting. It is well-understood that as speeds rise, Doppler frequency shifts increase, and this can stress receiver control loops. But in practical networking, increased speeds also bring new characteristics for fading due to the influence of speed, acceleration and motion on multipath and shadowing effects.

Comparing 4G/5G and Wi-Fi mobility

It is generally agreed that this is an area where 4G/5G is at an advantage, for a number of reasons. Following the classification above:

1) In the lower-frequency licensed bands, with high transmit power allowed, the effective range of a 4G/5G cell is of the order of ~2 km radius from the base station. A single Wi-Fi access point, mounted at elevation with clear line-of-sight can reach ~500 m - 1 km. For a given geography, more Wi-Fi access points than 4G/5G base stations will be required. [Note that this will change if a medium-band licensed frequency or an unlicensed frequency is used for the 4G/5G radio. In this case, there will be little difference.]

2) Inter-base station handovers with 4G/5G can be faster in a well-planned network, as there is a higher degree of client control from the base station.

3) 4G/5G as a system can operate with higher-speed clients.

Other considerations include very wide geographic areas. It would be expensive for a private company to build networks to cover public roads across a wide area, as it would have to acquire transmitter sites and arrange power and backhaul to them. But there are many examples of railway operators constructing dedicated Wi-Fi networks along their lines, using existing rights of way and power and fiber infrastructure to service base station access points. Similarly, oil field equipment including vehicular traffic can be serviced by a private, wide-area Wi-Fi network.

Finally, the maximum design mobility speed for the network is an important parameter. Wi-Fi is generally considered successful up to speeds of ~70 km/hr: beyond that, disruption at handovers can become significant. The IEEE 802.11 standards group has developed a variant, 802.11p, specifically for vehicular use and used a design goal of 200 km/hr, but general-purpose IoT clients do not use 802.11p today.

In contrast, today's LTE networks are considered to provide service to mobile clients at up to 350 km/hr – although anyone who makes cellular phone calls on the freeway will have experienced connection failures at highway speeds.

IDENTITY

A user or device presents an identity to the network as the first step for authentication. The form and flexibility of identities supported by the network is important. IoT devices are usually small and headless, so it is important that they use an authentication form that is physically small and preferably embedded in the device. Organizations managing private fleets of IoT devices need to acquire tools and build processes to program each device with a unique identity, keep track of the bindings between identity and device name/function/location, and set up an appropriate AAA (Authentication, Authorization and Accounting) server in their data center to handle authentication functions.

Identity in 4G/5G networks

The 3GPP mobile network uses the UICC SIM (Universal Integrated Circuit Card Subscriber Identification Module) card with its embedded identity, the IMSI (International Mobile Subscriber Identity) as the identifier for network access. This is part of the proprietary AKA (Authentication and Key Agreement) security framework used by 3GPP. The tamper-resistant SIM card stores other information, including encryption keys. SIM cards are now supplemented by eSIM (embedded SIM) modules, a form of software-programmable SIM that is embedded directly into a device. eSIMs are being incorporated in some new smartphones, and in IoT devices including connected cars.

As part of the 5G project, the 3GPP will broaden the available range of identity formats to allow more flexibility. The work is not part of the initial releases of 5G, but it is clear that support for non-AKA based authentication will include the EAP framework used by Wi-Fi (see below). Details are not final, but it is also clear that these new identity methods will be used for 'secondary authentication' and not be first-class options. For instance, they will not be available for roaming across public cellular networks.

An enterprise customer using a private 4G/5G network for IoT should ensure that sensors and other IoT devices and aggregators are SIM-capable or have eSIM functionality, and should acquire tools for programming SIM cards and and an AKA authentication server to authenticate them on the network.

Identity in Wi-Fi networks

While the cellular network relies on SIM cards and recently eSIMs, Wi-Fi has, since WPA2-enterprise authentication was introduced in 2004, supported a range of authentication types through the EAP (Extensible Authentication Protocol) protocol from the IETF (Internet Engineering Task Force). The primary identifiers supported by EAP are:

1) Username-Password based authentication through EAP-TTLS

2) X.509 certificate-based authentication through EAP-TLS

3) SIM, AKA and AKA' based authentication through EAP-SIM, EAP-AKA and EAP-AKA'

This means an end-customer can derive identity for secure network access through a number of methods, including username-password from the corporate directory (e.g., Microsoft AD), certificates from a private or public certificate authority, or the use of public or private SIM cards. Of these methods, username-password is considered the easiest to administer, followed by certificates then SIM cards. The first two options are easy to administer in an enterprise setting with existing infrastructure: identity stores, directory functions and AAA servers, while SIM authentication requires more complex software and processes.

Conclusion: Identity

In summary, the Wi-Fi network is very flexible with respect to identity, particularly for an organization managing its own identity system, while 4G/5G networks are much less flexible, relying exclusively today on SIM cards. Although the changes in 5G standards will bring them closer into line with the Wi-Fi framework over time, the alternate identities will not gain parity with AKA in the 3GPP architecture.

AUTHENTICATION AND ENCRYPTION

Authentication and encryption are the security counterparts to identity, and usually bound to the form of identity… but they are distinct functions in this context. While the identity of a device is an assigned label, uniquely applied to it, authentication is the process where the network assures itself of a new device's identity, discarding faked identities before proceeding with connection, and bidirectional authentication also requires the client to identify the network and verify that it is not connecting to a 'man-in-the-middle' or other malicious network device. Following authentication, connections and traffic flows are usually encrypted to maintain privacy. Encryption is particularly important in wireless networks, because monitoring of traffic is so easy – the eavesdropper needs no physical access, and can be some distance from the base station and client device. We always assume wireless communication can be monitored, so encryption methods are advanced. There is also an authorization function, where the network gives the device access to appropriate resources and privileges, but that does not concern us in this section.

Authentication and encryption in 4G/5G networks

The state of the art for 4G networks is AKA (Authentication and Key Agreement), a proprietary protocol of the 3GPP, which uses the SIM identity for the phone client and a special HSS (Home Subscriber Server) which includes an AuC (Authentication Center) function, located in the operator's core network to perform authentication.

4G implements mutual authentication, where the client device proves its identity to the network, and the network proves its identity to the client. This is done by taking the base identities previously provisioned in the SIM card and

HSS, hashing them to derive authentication vectors and exchanging these vectors over-the-air.

Encryption, or ciphering in 4G uses keys generated at the end of the authentication sequence. Encryption is symmetrical – it is applied in both directions. Different ciphering is used for signaling traffic and user-data. 4G (LTE) can use 128-bit or 256-bit keys for encryption, with 3 cipher options.

4G security is currently considered strong and unbroken. But as with all systems, vulnerabilities increase with time: the 3GPP recently deprecated SIM authentication, and 2G and 3G system security is considered breakable. Indeed, one of the most significant security threats in the cellular industry is from rogue base stations tricking client devices into downgrading to 2G or 3G connections.

3GPP security will be improved for 5G, with increased breadth from adding EAP options (see notes above). Enterprises adopting private 4G/5G systems should ensure they support state-of-the-art (4G at least) authentication and encryption for best security.

Authentication and encryption in Wi-Fi networks

Wi-Fi differs from the public cellular network, as it is used in many different settings and has consequently developed a ladder of authentication and encryption options, because increased security brings increased complexity and most consumers are sensitive to the amount of time and complexity required to configure their networks and client devices. The range of Wi-Fi authentication and encryption options covers:

1) Open access points. In some settings it is appropriate to have no authentication or encryption – some coffee shops, airports and other semi-public areas use open Wi-Fi with captive portal registration or click-through agreements, for ease of access.

2) Pre-shared keys. Most residential Wi-Fi access points are set up with a single pre-shared password. This allows new devices to be quickly added to the network. [Also known as WPA2-personal].

3) Individual keys. The fully-secure Wi-Fi security protocol is known as WPA2-enterprise. This uses the 802.1X framework where client devices have individual identity credentials, and a centralized authentication RADIUS server is used to authenticate them.

Industrial IoT installations should use WPA2-enterprise, as this is a fully-secure option. As noted above, the authentication framework uses 802.1X with EAP tunneling of authentication messages, where the authentication server authenticates the device, then authorizes the access point to allow it onto the network and generate encryption keys.

WPA2-enterprise can be configured with many different combinations of authentication and cipher algorithms and key lengths. The combinations used today in enterprises are considered secure.

[One of the few attacks on WPA2, known as KRACK, targeted a key-reinstallation vulnerability that existed in some access point and client implementations of WPA2. The Wi-Fi Alliance recently revised its tests to ensure that this vulnerability is no longer present in certified equipment… enterprise access point vendors have already released fixes in current software and client designers are following suit. The attack allowed decryption of client traffic and, in some cases, forging of spoofed traffic.]

Apart from the KRACK attack noted above, WPA2-enterprise is considered secure and certain combinations of algorithms and key lengths are certified for government secret-grade networks by the US and other governments.

The Wi-Fi Alliance recently announced new certifications that will provide improvements for each of the options above:

1) Open networks will provide encryption without authentication, so anonymous users can connect as before, but their traffic will be encrypted using OWE (Opportunistic Wireless Encryption) defined in IETF RFC 8110, a certification known as “Wi-Fi CERTIFIED Enhanced Open”.

2) Security certification WPA3-personal will supersede WPA2-personal (PSK) for home networks with pre-shared keys

3) WPA3-enterprise will supersede WPA2-enterprise. There is no new cryptography in WPA3-enterprise, it streamlines and simplifies the options available in WPA2-enterprise to provide two levels: 128-bit mode and 192-bit mode, where the 192-bit option is already certified by governments and is known in the US as NSA (National Security Agency) Suite B.

Conclusion: authentication and encryption

Both 4G cellular and Wi-Fi architectures support fully-secure authentication and encryption. When planning an IoT network, it is important to check that LTE AKA protocols, not older ones, are in use; while for Wi-Fi, WPA2-enterprise should be selected for best security. Although WPA2-personal with pre-shared keys may be considered adequate, as it has tradeoffs for simpler provisioning.

For example, 128-bit AES cryptography is an option for both 4G and Wi-Fi WPA2 networks.

Security researchers have discovered vulnerabilities in both Wi-Fi and cellular systems, but no practical, damaging attacks have been reported.

The levels of authentication and privacy of the two systems are equivalent: the key difference today is the broader range of authentication techniques available with Wi-Fi, although 5G equipment may catch up with EAP authentication options for private networks over the next several years.

REFERENCES

1) "5G Networks Role of Wi-Fi and Unlicensed Technologies", Wireless Broadband Alliance, https://www.wballiance.com/resources/wba-white-papers/

2) "Guide to LTE Security", National Institute of Standards and Technology, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-187.pdf

3) "LTE Security – How Good Is It?", National Institute of Standards and Technology, https://csrc.nist.gov/CSRC/media/Presentations/LTE-Security-How-Good-is-it/images-media/day2_research_200-250.pdf

4) "3GPP 5G Security", 3GPP, http://www.3gpp.org/news-events/3gpp-news/1975-sec_5g

5) "Wi-Fi Alliance WPA3 Specification", Wi-Fi Alliance, https://www.wi-fi.org/download.php?file=/sites/default/files/private/WPA3_Specification_v1.0.pdf

6) "Enterprises and multi-operator small cells", Small Cell Forum, http://scf.io/en/documents/069_-_Enterprises_and_multi-operator_small_cells.php

7) "The road to 5G: The inevitable growth of infrastructure cost", McKinsey & Company, https://www.mckinsey.com/industries/telecommunications/our-insights/the-road-to-5g-the-inevitable-growth-of-infrastructure-cost

8) "Investigation into the Doppler Component of the IEEE 802.11n Channel Model", Perahia et al, https://ieeexplore.ieee.org/document/5684207

9) "802.11p PAR", IEEE 802.11, http://www.ieee802.org/secmail/doc00270.doc

10) "Next Generation V2X SG PAR", IEEE 802.11, http://www.ieee802.org/11/Reports/ngvsg_update.htm

11) "LTE for vehicular networking: a survey", IEEE Journals, Araniti, http://ieeexplore.ieee.org/document/6515060

12) "Cellular Frequencies in the US", wikipedia, https://en.wikipedia.org/wiki/Cellular_frequencies_in_the_US

13) "Small Cell Products", ip.access, https://www.ipaccess.com/en/Small-Cells

14) "Cellular frequencies in the US", wikipedia, https://en.wikipedia.org/wiki/Cellular_frequencies_in_the_US

15) "Preparing 5G: take a very close look at IP router latency", Nokia, https://www.nokia.com/blog/preparing-5g-take-very-close-look-ip-router-latency/

16) "Policy and charging control architecture", 3GPP, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=810

17) "LTE; E-UTRA; UE radio transmission and reception", ETSI TS 36.101, https://www.etsi.org/deliver/etsi_ts/136100_136199/136101/14.05.00_60/ts_136101v140500p.pdf

18) "E-UTRA Physical layer procedures", 3GPP TS 36.213, http://www.3gpp.org/ftp//Specs/archive/36_series/36.213/

19) "LTE: find the iPhone that's right for your country or region", Apple, https://www.apple.com/iphone/LTE/

20) "Samsung Galaxy S9 Model Numbers and Country Variants", TeamAndroid, https://www.teamandroid.com/2018/02/27/samsung-galaxy-s9-model-numbers-variants/

21) "Containers in 5G and edge still under construction", AvidThink, https://www.fiercewireless.com/5g/containers-5g-and-edge-still-under-construction

22) "Will operators rise to meet the challenges of 5G?", Chetan Sharma, https://www.fiercewireless.com/wireless/industry-voices-sharma-industry-s-5g-moment-will-operators-rise-to-meet-challenge

23) 'Cradlepoint Wireless eyes CBRS…' , Fierce Wireless, https://www.fiercewireless.com/5g/apple-s-5g-iphone-may-slip-to-2021-not-likely-to-boost-sales-significantly

24) "3GPP 5G Security", 3GPP, https://www.3gpp.org/news-events/1975-sec_5g

25) "On cellular encryption", Matthew Green blog, https://blog.cryptographyengineering.com/2013/05/14/a-few-thoughts-on-cellular-encryption/

26) "With all the attention to 5G, don't ignore Wi-Fi", Mark Lowenstein, https://www.fiercewireless.com/wireless/industry-voices-lowenstein-all-attention-to-5g-don-t-ignore-wi-fi

27) "Apple's 5G iPhone may slip to 2021…", Fierce Wireless, https://www.fiercewireless.com/5g/apple-s-5g-iphone-may-slip-to-2021-not-likely-to-boost-sales-significantly

aruba

a Hewlett Packard
Enterprise company

Contact Us Share