# What is 5G? And do we need it?

*Being a brief history of influence in the mobile industry.*

The mobile industry loves hype. Now that 4G phones have reached the market, suppliers are keen to promote the next dollop of "jam tomorrow" by offering the world 5G - something that's still rather nebulous, but as always in this industry, allegedly better than what we have today. Most users have still to experience 4G, but that's par for the course. The industry loves something new, preferably with a bigger number. It begs the question of whether we need it, and even what it is? To try and answer these questions it's instructive to look back at the history of mobile to see just what the "G"s mean.

In the beginning was 1G, but nobody called it that. It was the formative world of analogue mobile phones, commonly known as bricks. It was named 1G retrospectively, long after digital phone systems had come to market. In those early days of digital, if you'd asked anyone in the industry what "G" stood for they'd probably have said GSM – originally a European standard called Groupe Spéciale Mobile, but quickly anglicised to Global System for Mobile Communications when the industry realised that people who didn't speak French might be interested in buying it. GSM evolved out of an EU mandate which decreed that Europe should have a single, interoperable digital mobile phone standard. Although that was partly to give an advantage to European manufacturers, it was a far-sighted decision that brought interoperability to what had been a fragmented mobile market, where business travellers might need half a dozen different phones to stay in touch. It took the US a further ten years to understand the benefits of common infrastructure, but that's another story.

The original GSM standard, which is now referred to as 2G, stood for Second Generation. It was a new digital approach to mobile telephony which brought in a new audio clarity for voice calls and enabled data up to a maximum of 9.6kbps, later increased to a blistering 57.6kbps, although few networks or phones supported more than 28.8kbps. (Bear in mind that landline dial-up modems in those days started at 2.4kbps, eventually evolving to about 40kps, the latter being cynically marketed as 56k. So at this point the 2G data performance was pretty good.) The 2G standard heralded the start of mobile data, although many applications eschewed the data capabilities of 2G and used SMS instead.

The promise of mobile data fascinated the networks. Voice and SMS revenues were starting to rise nicely, but data gave the opportunity of selling contracts for connected machines – the start of M2M (machine–to-machine) communications. The prospect of regular, recurring revenue from devices which would never call customer support was too great a prospect for the industry to resist. However, GSM data had a problem – it was circuit switched, which meant that like dial-up modems you needed to establish a phone call before you could send data. That didn't chime with the way the internet worked, and how M2M devices preferred to use a data network. They just wanted to send packets of data. So the clever chaps who wrote the GSM specifications came up with a new version which included a feature called Generalised Packet Radio System or GPRS, which brought the Internet model of packet connectivity to mobile phones. Fondly known as 2.5G, GPRS didn't go much faster, having a maximum data rate of 114kbps. Few networks supported rates beyond 56kbps, but for M2M that didn't matter, as most devices only sent a tiny amount of information. Because it was a packet data approach, data could be sent quickly and cheaply with GPRS. As a result, cellular modems for machines began to appear.

GPRS didn't result in many M2M contracts being sold, but it made small data transfers faster. As a result we got the first simple Internet browsers and applications on smart phones. For phone users the GPRS data plans were expensive and it still didn't give them much of an Internet experience, so the main consumer applications were for two very personal services that users were traditionally prepared to pay for – Girls and Gambling. The practical meaning of those two "G"s had become clear.

Even then, most revenue still came from voice, as networks took advantage of mankind's, and increasingly woman and teenagekind's innate desire to talk to one another. It became apparent that the way that the 2G network operated meant it would have difficulty in coping with the demand, so work began on a new, third generation of the standard – 3G, which would cope with the increasing number of users. And it would allow increased data throughput, pushing usable data rates up to 100kbs and beyond.

At this point Governments around the world threw a spanner into the works. They saw the amount of money that mobile operators were making and decided they'd like some of it for themselves. They'd not made a vast amount of money out of licensing the spectrum for GSM networks, as no-one had been sure that the gamble of building a new digital infrastructure would pay off. Now that everyone could see it had, Governments saw the opportunity to auction the spectrum for 3G to the highest bidders, often making it clear that they would allow new mobile operators to join in, with the intent of ensuring a bidding war.

It led to a period of collective madness, where operators fought to outbid each other. To justify their bids, armies of consultants dreamed up ever more exotic applications for 3G, with the hope of increasing ARPU. To that end, numerous speculative new data applications were proposed, the most memorable of which were video calling and Multi Media Messaging. When the dust settled and the Government coffers were filled, 3G users decided to ignore these and carry on doing more or less the same as they had with 2G. The only difference was that the more efficient data network allowed more downloads and better latency for connected applications. 3G became Girls, Gambling and Games.

Over the next few years, data rates improved to the point where the 3G networks could justifiably claim they were a mobile broadband experience. For the first time, users could get something close to an Internet experience on their phones. Operators, notably Cellnet in the UK, had tried to pre-empt this using GPRS when they rolled out WAP in 2002 (commonly known as the Worthless Application Protocol or Wait and Pay because of its sluggish response). The experiment had been less than satisfying. As the infrastructure evolved through EDGE and HSDPA to support higher data rates, that restriction disappeared. Now the only limitation to a decent mobile Internet experience was the phone, as none had yet been designed with internet usability in mind.

Then came the iPhone. To be accurate, we ought to call it the iPocketComputer. Compared to previous phones, where smartness was generally little more than a few geeky applications tacked onto a voice phone, the iPhone's ease of use was a revelation to consumers. They fell in love with it and its applications, suddenly consuming more data than the network operators had ever imagined. Sadly for the operators, consumers weren't prepared to pay substantially more for it. In a disruptive move, users started to use Wi-Fi networks to support more of their data access, depriving the network operators of their expected revenue. The faster Wi-Fi data rates began to make mobile video a more compelling experience, which led to market pressure to replicate it when using 3G. To help keep hold of their customers, operators responded by offering "all you can eat" data plans along with the promise of higher speeds. The result was that the

network infrastructure once again started to reach capacity, beginning a vicious cycle of investment to maintain the data capacity to cope with customer demand.

Incidentally, the use of Wi-Fi to offload mobile data wasn't new. As far back as 2004 the industry was working on UMA – Unlicensed Mobile Access – a method to allow phones to seamlessly roam from mobile to Wi-Fi networks. Network operators and new Wi-Fi cloud businesses saw this as a way of getting more out of their current infrastructure. A few companies even offered the service for corporate users. But it needed the iPhone's Apps to stimulate wholesale consumer demand. Instead of being a corporate convenience that the network operators could charge for, it became a practical necessity to prevent their networks from being overwhelmed.

Although the network operators had mostly overpaid for their 3G licences, if nothing else had changed they could now have contemplated sitting back and reaping the profits from their infrastructure for the next decade. But the balance of power was changing. With the launch of the iPhone, Apple had changed the pecking order. From this point on phone vendors took over the reins of defining the future direction of the mobile industry.
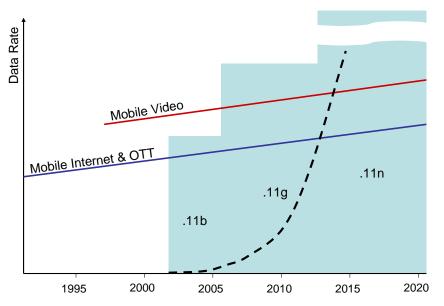
As if this were not enough disruption, something else nasty was happening to the operators. With a robust mobile data network now in place, application developers started building what were termed "Over The Top" or OTT applications. These are services like Skype, WhatsAPP and Snapchat which replicate traditional voice and messaging services using the data network and playing to the new "all you can eat" data plans. They were either free or lower cost than the equivalent network offering, cannibalising still more of the ARPU, and consuming even more data.

It meant that once again operators needed something that would provide even more capacity. The boffins from the standards group which had now renamed itself 3GPP – the 3$^{rd}$ Generation Partnership Project, came up with a new acronym – LTE or Long Term Evolution. It was yet another version of the standard which offered even more data. For the first time, in a sign of the way that mobile telephony was evolving from mobile everywhere to data everywhere this new standard did not even support voice – it was a data only infrastructure. That became an issue when some operators realised that they might need to turn off their 2G and 3G networks to get enough spectrum for LTE, at which point Voice over LTE or VoLTE was developed. Ironically it's not dissimilar to the OTT voice services which prompted the original development of LTE.

One of the things which was causing yet another step-rise in data usage was social networks. Users became addicted to this new form of digital gossip. Their phones' inbuilt cameras made it ever easier to upload photos and the selfie was born. Privacy became as outdated as the twitching curtain. Consumers had now defined 4G without any help from the network operators. It was Girls, Gambling, Games and Gossip.

This highlights another aspect of the change in power within the industry, with phone manufacturers taking decisions away from network operators. For many years the network operators bought the vast majority of phones and sold them on to consumers. The operators strongly resisted any feature going into the phone which didn't make them money. They couldn't see the point of GPS, sensors or cameras. They eventually relented on cameras when they tried to promote video calls using 3G, only to find that users had a totally different idea about how to use them. Whereas the operators used to control the specification of the phones they sold, today those decisions are made by phone manufacturers who now design them for customer appeal, not for network operators.

We see the same change in the balance of power towards phone manufacturers and away from network operators with the incorporation of Wi-Fi into handsets. Before 2005, virtually no phones incorporated Wi-Fi chips and operators resisted the additional cost. As internet use grew that changed rapidly, with phones using Wi-Fi to support data rates of multiple Mbps. Today almost every smartphone has Wi-Fi capability – a feature driven almost entirely by user demand.

**The pressure of the Wi-Fi experience as an Alternative Carrier**



In many ways, the operators only have themselves to blame. From texting to selfies, they have consistently made the wrong call on what consumers want to do with their phones. With more and more Over The Top data applications appearing, what customers do with their phones is now totally out of the operator's hands – all they can do is suffer the extra bandwidth requirements and look back on the days when they had hopes of controlling walled gardens, rather than letting their customers control them.

At this point it's worth looking at the politics of this evolution, as the power base has shifted and is shifting again. In the early days of GSM there were three players: the companies making mobile phones, the network operators and the companies making the base stations and infrastructure to support the mobile networks. All three were interdependent on each other but power was firmly in the hands of the networks who were the paymasters for the other two. With the boom in consumer mobile still five years away, the operators paid the bills and dictated the technology. This was the age of the start of the mobile industry, largely European centric, which was separate and very different to the US centric PC industry. It was ruled by service provision and monthly contracts, not by hardware
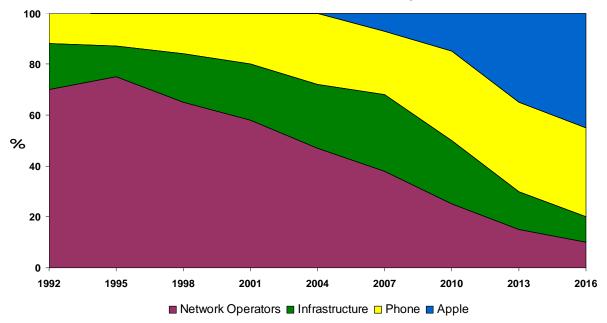
As the industry continued to grow, the balance of power started to shift. Growth favoured a few large mobile phone manufacturers who benefitted from economies of scale. At its peak, Nokia produced around 40% of all GSM handsets. Mobile operators also consolidated, with the successful ones becoming global companies. Infrastructure companies followed the trend whilst the PC industry looked jealously at the market size and began to hanker for more of the opportunity.

As consumer demand grew and the Chinese market opened up prices were forced down, not only for phones, but also for infrastructure as the future giants began

to emerge; first Samsung, LG and HTC in Taiwan, then Huawei, ZTE and Xiaomi in China. They not only made things more cheaply, they also challenged the technical competence of the Western incumbents, starting to become major players in standards development. The established hierarchy of influence amongst mobile companies began to splinter.

From its initial point of power, the influence of mobile operators grew as they consolidated, since when they have been consistently squeezed. Infrastructure providers used 3G to flex their muscles, taking advantage of the new, emerging data specifications as the mobile internet took hold. As we're about to see, that level of influence is now waning as they have evolved technology too fast. The winner has been the new generation of smartphone vendors, providing consumers with a product that is more compelling than any service offering. It is telling that most operator branded stores now devote more space to promoting phone brands than their own.

Apple applied the killer punch with the introduction of the iPhone. You cannot overstate the importance of it being a computer in your pocket, as everything that followed on from it has been about the Internet and Data, not about voice. For the first time, the dynamics of the PC industry have begun to dictate the course of mobile evolution, although the traditional PC industry players have still not necessarily benefited from that influence. The Apps store, probably even more than its creators ever envisaged, drove the evolution of OTT services; with them the PC spawned Wi-Fi infrastructure became a permanent part of most user's mobile phone experience.
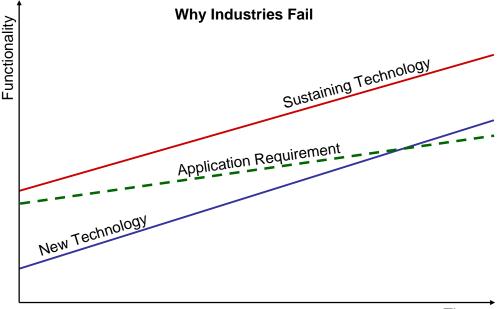
## Influence in the Mobile Industry



The chart above illustrates the change in influence within the industry over its quarter century existence.  The irony is that many of those who were responsible for its early years of growth still fail to realise that they are no longer leading, rather than being led.  The original pioneers of the mobile phone experience – Nokia, Siemens, Philips and Ericsson are gone. Motorola soldiers on in name alone, having been sold to a Chinese laptop manufacturer and Blackberry is as bitter as its namesake after Michaelmas. The Devil has comprehensively urinated over the founders of the mobile handset business. Whether the Nokia or Motorola names will continue to adorn any phone after 2015 looks increasingly unlikely.

Network operators hang on, behaving more as utilities with brand pretensions than the proud companies they once thought they were. Apple's recent introduction of the nanoSIM may be the final straw that consigns them to utility status. It allows a phone or tablet to be sold pre-provisioned with a choice of network operator, letting the user select their network provider from an application menu, devoid of any operator marketing input. If that approach becomes standard, then the networks really are no more than a pipe.

Which leaves the infrastructure providers. The natural course of most infrastructure operators (think gas, electricity, water, etc.) is to deploy and then maintain the system, working to extract as much revenue as possible from their investment whilst minimising any future spend on it. They have little interest in upgrading it, rather than milking it. That's not a good model if you're a fast growing company making base-stations. There's no doubt that the world will continue to want more base stations for many years to come, but it would be so much better for an infrastructure manufacturer's balance sheet if they could persuade their existing network operator customers to upgrade what they already have. Hence their continuing push for something even shinier with more flashing lights, all justified by what they claim is the consumer's insatiable desire for more bandwidth and faster download speeds. That's the model they managed to sustain from the start of 2G to the 4G networks we have today. They'd like to continue with that model, hence the talking up of 5G.

Before looking at what 5G is, I'd like to refer back to a classic book about business cycles in the technology industry – Clayton Christiansen's "The Innovator's Dilemma". In it he looks at why technology companies fail. Often it's not because existing industries fail to innovate – in fact they do the opposite – they innovate the sustaining technology at too fast a rate, beyond what the market needs. If customer doesn't need the level of performance that innovation brings, then a cheaper technology that had previously been inadequate, but which is also innovating fast may provide a lower cost alternative, making the incumbent uncompetitive. Effectively the status quo becomes too complex for its own good. The principle has been seen across many industries, from disk drives to mining equipment.



**Why Industries Fail**

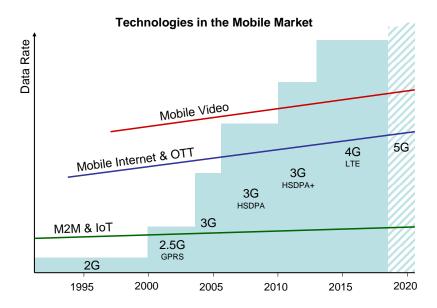*After "The Innovator's Dilemma" – Clayton Christensen*

The iPhone is an interesting exception to this rule. That won, not because it followed this pattern, but because it replaced the mobile phone with something else – a highly usable pocket computer which could do phone calls. It couldn't exist until there was sufficient data capability within the networks, but its data needs, along with its rivals, are no longer being constrained by the infrastructure. With this new usage model, smartphones have developed a startling uniformity. To most users today, if their phone is not a slab of black glass, it's not a phone. That's an interesting challenge for Apple - how long can they keep their position before commoditisation kicks in and they become the next participant in the mobile ecosystem to lose their power base?

What does all of this mean for those espousing 5G?. If we put some numbers into the sustaining technology curve we can see some interesting effects. The figure below looks at the application requirements to enable a usable experience for three key applications – M2M, a consumer Internet experience (which should be as equally compelling as their PC based experience) and mobile video streaming. Underneath these we can see the typical speeds of the various GSM infrastructure deployments. It makes it obvious why many of the early mobile Internet and video services failed – there wasn't enough data throughput to make them compelling consumer applications.

It wasn't until 2006 that networks could provide the infrastructure to meet the requirements for a useable mobile internet experience. It's why smartphones didn't take off before 2007, which is when the iPhone arrived. Not only did the iPhone have a far simpler user interface, but it appeared at the point where the networks could provide the data speeds to support the explosive growth of applications.

Similarly we can see why mobile video on demand has only grown in the last few years – it needed that extra amount of throughput that came with HSDPA. Looking forward, the data requirements for both mobile internet and video will continue to rise steadily, as video formats and phone and tablet screens support higher definition playback, and because application developers will continue to generate more bloated applications. However, there's nothing on the horizon that won't be served by current 4G technology. 5G speeds for the smartphone market are not needed. It's the classic trap of providing rates of performance improvement that exceed the performance the market needs, which often heralds a point of major disruption for the incumbent suppliers.
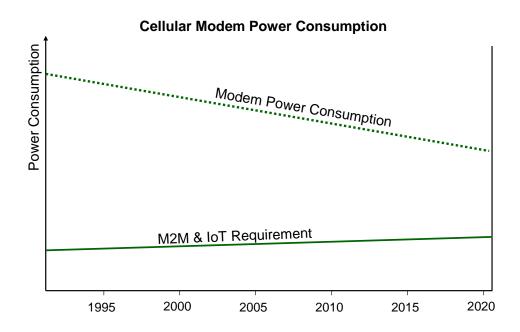
Which brings us back to the question of what 5G is and why we need it? 4G has its own evolution path, driven by competition between infrastructure companies.

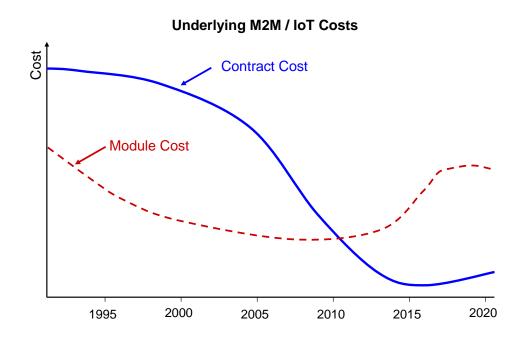**Technologies in the Mobile Market**

The current plans for 4G envisage peak data rates that could reach 1Gb/s, way ahead of any perceived mobile user need. Perhaps aware that speed may not be the selling point it once was, others have suggested that 5G means more efficiency, lowering the cost of transporting data for operators, by cynically selling them more expensive equipment. Others look to dynamic spectrum usage or moving up to the 20GHz bands and higher. More perceptive suppliers who have listened to market sentiment have started marketing the interim step of 4.5G, not because it's necessarily needed, but because it sounds less speculative than 5G. ZTE have gone a bit further and offered "pre-5G", whatever that may be.

It has been pointed out that a more honest meaning for LTE is Long Term Employment, as the 3GPP bandwagon sponsors thousands of engineers to travel around the world to meetings and testing events in support of the emerging standards. So maybe our fifth "G" should be Gallivanting. The industry could save a lot of money if it recognised that what we have is adequate.

The M2M requirement in the figure above also merits examination. Most M2M and Internet of Things applications transfer very small amounts of data – so small that early GPRS networks could easily cope with them. If we look back at the graph, M2M has not been held back by data throughput. The lack of growth is due to other constraints, largely dominated by the cost of hardware and data contracts, but also by the power requirements of any GSM technology that been developed so far.

**Cellular Modem Power Consumption**



While contract costs are coming down (I've seen them as low as $0.50 / month), hardware is still around $7 in volume for the simplest GPRS modem. As GPRS networks get turned off to make more room for LTE, the more complex 4G modems will triple the current price points, making a cellular solution for M2M less competitive.

## Underlying M2M / IoT Costs

Cost

Contract Cost

Module Cost

1995    2000    2005    2010    2015    2020

The 3GPP standards group have a couple of proposals to try and address these issues – shaving a bit off the cost with Category 0 of LTE Release 12, which promises "low cost and enhanced coverage" and the more promising LTE-M forum report which suggests a clean slate approach supported by Huawei and Neul (or more accurately Huawei and Huawei, as they've just acquired Neul). The most interesting part of this is that it is a true clean-slate approach and hasn't anything to do with LTE, but is being developed by the GERAN group, which still looks at basic GPRS developments.

I've pointed out that we desperately need a new Low Power Wide Area Network like this if the Internet of Things is going to take off. The next ten billion connections won't come from people, but the things around us which we want to connect. Forget Gallivanting. Forget even Girls, Games, Gambling and Gossip – they're just sustaining "G"s. Instead we need a solution for the everyday things like Golf clubs, Garden gnomes, Garage door openers, Gym equipment and Garbage bins. What the industry and the network providers need is something new that lets them start making money from the billions of thing around us. It gives them the opportunity to wrest back their sphere of influence from the smartphone manufacturers, who are heading for their own cliff of commoditisation. It's time the industry took a leaf out of the Innovator's Dilemma and realised that it can walk away from the never-ending and expensive race for something faster purely for its own sake. Users don't need the 5G that is being pushed by the momentum of Long Term Employment. Nor do operators. The future no longer lies in users, but the things they use beyond the handset.

## Nick Hunn

November 2014


nick@nickhunn.com

+44 7768 890 148


Read more on my Creative Connectivity blog.