



THE ART AND SCIENCE OF ANTENNAS FOR CELLULAR SERVICES AND THE IOT

New devices and spectrum are challenging antenna design – Brian Collins explains

ANTENNA DESIGN for many familiar applications remains something of a dark art to most electronics engineers. The last couple of decades has seen an explosion in the number of devices that employ radio communications and the number of systems and frequency bands that support them.

This article describes some of the challenges and the antenna techniques that are used to meet the often conflicting requirements of efficiency, bandwidth and fashionable industrial design.

WHERE WE STARTED

Many early mobile handsets were fitted with various kinds of pull-up or swivel





antennas that had to be deployed to make a phone call. These were inconvenient to use and often easily broken. They shrank to small, finger-joint-size encapsulated helical antennas. Developed versions of these were made to cover dual operating frequency bands (900/1800MHz in Europe and many other countries, 850/1900MHz in the Americas, and some other bands in East Asia).

It was not until the coming of 3G that truly worldwide roaming could be supported by just two air interfaces (GSM and UMTS) and five frequency bands. The antenna designer usually designed these penta-band antennas to cover two broad bands, 824–960 MHz and 1710–2170 MHz. The antenna is standards agnostic and the different

band combinations used in most of the world could be accommodated with a single penta-band antenna design. A typical example of the most common design (the IFA or PIFA) is shown in Fig. 1, developed in its dual band form by Liu¹.

INCREASING COMPLEXITY

From the mid 90s, handset designers began to add features to what had earlier been just a mobile telephone, but surprisingly when 3G was introduced no one knew what its “killer app” would be.

By 2000 many handsets were fitted with cameras, and they started to incorporate larger screens, music players and Bluetooth, all needing more power and bigger batteries. Together with the trend for smaller and thinner handsets, this created intense competition for space in the product.

At the same time the industrial designer wanted external antennas to vanish and so internal antennas were required. Antenna designers had to be able to produce usable performance from an ever-diminishing volume, shared with cameras, loudspeakers, and noisy processors and memory operating at ever faster speeds.

During this era the performance of handset antennas got worse. Hundreds of designs flowed from scores of design studios, who often poorly understood the compromises of space, antenna position, platform configuration and intrinsically noisy circuits.

While antenna designers knew that it was impossible to get high efficiency from a small antenna in a small platform, both in practice and in accordance with theory, the industry seduced users to ‘upgrade’ their handsets with smaller, sexier replacements offering the latest gizmos, but radio performance that was often much worse than the early, larger handsets, especially ones with external antennas.

BASE STATION ANTENNAS

The main evolution in network practice in the mid 90s was a widespread move from space diversity at base stations to polarization diversity. A single antenna package now contained two antennas

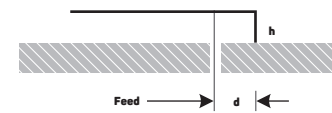


Fig. 1(a)

The basic configuration of an IFA/PIFA

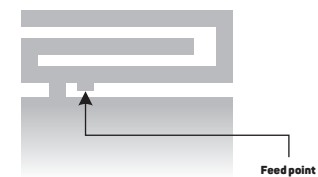


Fig. 1(b)

A typical dual band PIFA

operating with orthogonal (isolated) slant linear polarization at $+45^\circ$ and -45° . This not only reduced the cost of base stations (one antenna instead of two separate ones), but reduced their visual profile, allowing operators to densify their networks while attracting less criticism of their visual profile.

Single-pole and streetworks base stations became common sights. Further developments around 2000 enabled a single physical antenna package to support dual-polar operation simultaneously on low band (850/900MHz) and high band (1800/1900/2100MHz).

THE SMARTPHONE

Development of the handset as a mobile entertainment device took a decisive step forward with the launch of the original iPhone in 2007. This popularised a much larger device format with a large screen and also enabled improved radio performance.

The link between platform size and bandwidth was noted by Vainikainen². With an unbalanced antenna, the platform PCB carries RF currents. These currents will radiate effectively, irrespective of the dimensions of the device used to excite them – the thing we usually describe as being the antenna. These radiating currents must be carried by the internal





PCB if the case is plastic, or on the outside of the case if it is made from conductive material such as metal.

LATER 3G

The later phases of 3G development introduced higher data rates, supported by high speed packet access (HSPA). To improve the reliability of data transmission, smartphones introduced diversity in the handset. This required two multi-band antennas having reasonably comparable performance and providing uncorrelated signals, meaning that the antennas had to be reasonably isolated from each other.

The pressure on space was now intense. The addition of GPS, two cameras, dual-band WiFi, Bluetooth, larger batteries (to support ever faster processors and large screens) resulted in the increasingly tight integration of



SAMSUNG S100

Dipoles were popular in dual band phones but as RF engineers lost power in the networks purchasing fashion has given way to function

antennas with the rest of the hardware.

By the late 2000s it was obvious that the smartphone was not a passing trend. The utility of a pocket-size computer with access to the web and e-mail was evident to all. The use of real-time streaming video became well established and new solutions were needed to support ever-growing traffic.

THE COMING OF LTE

Planning for 4th-generation mobile

networks mandated a huge increase in spectrum for mobile operators. This resulted in the (still-growing) list of more than 40 frequency bands, allocated in a disorderly patchwork across the world. The economics of mass production require product standardisation, so manufacturers need a common solution that can be used (almost) anywhere in the world.

The result of this pressure was the advent of small antennas that can be tuned under microprocessor control. Relaxing the requirement for instantaneous bandwidth and replacing it with frequency-agile narrower bandwidth addresses the size to bandwidth trade-off, but increases the degree of integration of the handset antenna, which now needs not only mechanical and electromagnetic integration, but also adaptive tuning





controlled by the platform.

As a result of this ever-increasing integration, major smartphone vendors took antenna design in-house and many independent antenna design firms and consultancies looked for new markets.

MIMO (multiple-input multiple output) is generally thought to be a 5G technology, but the increase in spectral efficiency available by using multiple antennas and clever signal processing had been known for some time, with later flavours of the 3G HSPA technologies and now 4G LTE, the use of MIMO has become firmly established. This uses multiple antennas both at the base station and at the user equipment. MIMO works best when the signal-to-noise ratio is high, so adding more antennas while allowing their performance to be degraded is no solution.

LATER DEVELOPMENTS: LTE/A

LTE-Advanced further tests the ingenuity of both antenna and platform design engineers by using channel aggregation. This allows a single, high-speed, data stream to be transmitted by combining capacity on multiple frequency bands, something easy in principle using OFDM, but needing antennas that can operate concurrently on multiple frequency bands, now spread over as much as four octaves of spectrum.

While later releases of LTE/A enable the aggregation of larger numbers of channels and bands, it is not currently clear where this may lead smartphone development.

Short latency is a requirement for many games, but exceptional data rates are less often needed. And how can operation over multiple bands be reconciled with antenna tuning?

However, the widespread use of smartphones as a link from laptops to the mobile network means that it is desirable to be able to support high-definition video. While ever-higher orders of MIMO are planned (8x8 MIMO using 256-QAM modulation), smartphones may not be their real market.

An earlier trend to providing mobile network connectivity to laptops seems to have slowed, but as a large platform like a laptop could support high-order MIMO (with many isolated antennas) a new application could stimulate this market.

5G

The extension of frequency bands for 5G mobile applications includes millimetre-wave bands around 28, 52 and 70GHz. While these may not be used for day-to-day communication with smartphones, they could be used at crowded public venues like sports stadiums.

Antennas for these frequencies are likely to be in the form of smart arrays providing adaptive radiation patterns. The size of antenna elements ceases to be a problem and smart antenna arrays can be integrated with the radio chips. (Testing the products though introduces a whole new set of issues!)

GETTING THE BEST FROM AN ANTENNA ON A SMALL PLATFORM

Provide continuous ground planes on both outer faces of any PCB that carries an antenna, stitching them together at their edges with closely spaced vias.

Mount typical IFA/PIFA antennas at one end of a small platform, allowing the excitation of radiating currents along the longest dimension of the PCB.

The optimum antenna may be a loop or notch, located near the middle of the longest side of the platform.

Remember that the radiation pattern of any antenna on a small PCB will be the characteristic donut-shape of a dipole with its hole aligned with the long axis or the diagonal of the PCB; make sure that the orientation in use will provide radiation in the wanted plane.

Any antenna will need matching in place on the complete device in working configuration. Failing to do this may reduce the range of the device by a factor of two or more.

Any antenna works to some extent, but if you are concerned about maximum range or long battery life sub-standard performance is not good enough.

Most plastic cases are transparent to radio waves but have a marginal detuning effect on internal antennas, easily compensated by a matching circuit. Metal cases are likely to need special antenna design.

THE INTERNET OF EVERYTHING?

The earliest application of cellular networks to data communications with “things” was circuit-switched data (CSD), an early feature of GSM that provided up to 9.6kb/s. The costs of the hardware and connection services were high, and applications were limited to simple remote alarms, monitoring and control. SMS was also adopted when only small amounts of data had to be carried.

Early modems were complete packages, often fitted with coaxial connectors for use with external antennas. These were often simple quarter-wave monopoles or dipoles that were placed on top of



INTERNET OF BIKE LOCKS

The first generation of the Biseccu bike lock only has Bluetooth connectivity but future generations, being planned now will use techniques outlined in this article to develop a cellular solution.



STREET LIFE

Better antenna design in devices allows base stations to be more elegant.

an equipment rack or close to an external window. Described as machine-to-machine (M2M) communications, both ends were usually static.

The increased data rates possible with 2G (HSCSD, GPRS and then EDGE) and the integration of mobile data with the internet, as well as the falling costs of hardware and service, encouraged much wider use of M2M (and we started to call it the IoT). Modems were now widely available as integrated packages, and the lower cost of 2G modems over a long period has led to a large installed base of 2G equipment. Even now, 2G/3G modems are significantly less expensive than 4G, so again there is a gap between what is possible (and is used in entertainment and social media applications) and what is most economic for IoT that needs low cost.

A wide range of commercial antennas for IoT devices is now available for use in single or multiple frequency bands. These mostly comprise dipoles, monopoles, loops or PIFAs. The best are very competent designs, which if well

implemented on a user platform will give really good results. Regrettably some offerings have poor or indifferent performance and many developers are unable to measure their performance.

The functioning of the antenna is always strongly dependent on the interaction between the antenna and the platform. Obtaining the best

performance relies heavily on the platform environment, the suitability of the radiation pattern to the application and on the implementation of a matching network that is always platform-specific.

ROLE FOR REGULATORS

An opportunity for system optimisation has been missed in today's assignment of mobile network frequencies for low power devices. Given the relationship between efficiency and bandwidth, it would be desirable that the frequencies assigned for these devices by every network were grouped closely together. This would allow small devices to provide high efficiency albeit over a small bandwidth.

The opportunity has so far been lost, with widely scattered frequency allocations, so for reasons of minimising production cost, a device requiring only a small operating bandwidth must be capable of operating in multiple allocated bands.

THE PLATFORM ENVIRONMENT

Adding connectivity to an existing platform – for example a gas or electricity meter – is never a simple matter. An electricity meter is subject to severe mains-borne noise, while a gas meter must not have mains power connected. Both may be mounted outside, but many meters are located deep inside buildings. Both require only occasional communication, perhaps polled from a remote controller. Domestic water meters are usually mounted underground and again have no mains power. These all have challenging link budgets, yet reliable connectivity and minimal energy consumption are essential.

As the possibilities of connected devices become increasingly recognised,

sensors, monitors and personal aids are being created that need reliable network connections. Many of these need to be small and light, while others have no suitable location for a radio device. As an example, think of adding sensors and radio communications to a beer keg: potentially useful, but not easy!

Large platforms create issues of obtaining reasonably omnidirectional coverage, and antennas on small platforms have the intrinsic problems of limited bandwidth and efficiency.

SOLUTIONS

New air interfaces such as NB-IoT and LoRa aim to have better link budget than cellular and support long battery life, but these improvements can easily be negated by a poor antenna.

It is common wisdom that free space propagation loss is lower as frequency is reduced, though the reason for this often misunderstood. If a device radiates 1 watt of RF energy equally in all directions, the power density d metres away is simply given by 1 watt divided by the area of a sphere of radius d , independent of frequency.

To find the power received by another antenna at distance d away we need to know its effective area. Not surprisingly, the effective area is related to the physical size of the antenna – for a half-wave dipole it is 0.41 times the square of the wavelength. This means that, for an antenna (or more usually a platform) of fixed physical size, the power that it is possible for it to receive in principle remains constant as frequency is reduced. But, if the antenna keeps its electrical size (which we usually want it too both for easy impedance matching and best omni-directionality), its effective area reduces as the square of frequency, and so does the received signal power.

Unsurprisingly, the efficiency of an antenna falls as its length is reduced relative to a half-wave dipole. Reducing the frequency has the benefits of reducing losses caused by diffraction and penetration through walls, but if the transmitting and receiving antennas are fixed in dimensions there is no other benefit, and the efficiency of an electrically small platform will fall as the

MEASURING PERFORMANCE

The performance of a radio platform is characterised by the following parameters, which should be measured across the working frequency band(s):

a) The input impedance of the antenna is directly related to reflection loss, caused by energy being reflected back to the transmitter instead of being radiated. It is measured by connecting a network analyser to the input of the antenna matching circuit.

b) Radiation patterns show how radiated energy is distributed in 3-D space around the device. It is measured in free space or an anechoic test chamber, by sampling the radiated field in all directions.

c) Efficiency is the ratio of the total radiated power to the power delivered to the antenna by the transmitter. This is measured at the same time as the radiation patterns but requires absolute calibration of the chamber to determine the power flux in every direction in space.

d) Total radiated power (TRP) measured in a calibrated chamber is the integral of the power density over a surface completely surrounding the device. Typically it is approximately equal to the power delivered by the PA multiplied by the efficiency of the device.

e) Total isotropic sensitivity (TIS) is the receiver sensitivity averaged over a surface surrounding the device. This measurement requires a loop-back from the DUT to determine the signal needed for a defined minimum BER to be achieved.

frequency is reduced. Also, the antennas tend to become more directional which does not make installation easy.

An electrically small antenna (or an antenna on an electrically small platform) will always have a narrow operating bandwidth, and attempts to increase its bandwidth will result in lower average efficiency. For example, a small platform operating in the very narrow 868MHz ISM band will have a higher mean efficiency that of one of similar size operating over the whole band 890-960MHz.

An opportunity for system optimisation has been missed in today's assignment of mobile network frequencies for low power devices. Given

the relationship between efficiency and bandwidth, it would be desirable that the frequencies assigned for these devices by every network were grouped closely together. This would allow small devices to provide high efficiency albeit over a small bandwidth.

The opportunity has so far been lost, with widely scattered frequency allocations, so for reasons of minimising production cost, a device requiring only a small operating bandwidth must be capable of operating in multiple allocated bands.

Many small devices have batteries that occupy a large proportion of the available internal volume, often limiting the scope for providing efficient antennas. The outer conductive case of a lithium rechargeable cell is usually positive, so with a conventional negative ground on the device circuit there is likely to be a significant RF voltage across the battery terminals. Unless a low-impedance path is provided across the battery this will lead to the loss of precious RF power in the battery and also being fed into the battery line of the device where it will be dissipated in internal circuits.

IN CONCLUSION

In any electrically small device, RF performance is a function of the whole platform; the 'antenna' is only part of the story. Antenna engineering is not a black art, as sometimes imagined, but is an engineering skill essential for modern high-performance radio devices. **CWJ**

BRIAN COLLINS RF GURU

Brian Collins has spent his career in the design and manufacture of antennas for communications and broadcasting. After a long career with CSA Ltd (now part of Amphenol) and spending 10 years with Antenova, he now operates a consultancy designing antennas for IoT, comms and consumer devices. An honorary visiting professor at Queen Mary, University of London, he is a champion of CW's Radio Technology SIG.

Continue the conversation on Twitter. Follow us on @cwjpress hashtags #cwjournal and #antenna

