RTT TECHNOLOGY TOPIC
July 2015

5G Modulation and Coding

Last month we discussed how reducing backhaul delay and delay variability improves scheduler efficiency and how and why this is beneficial to 5G mobile broadband network economics.

We suggested that scheduler efficiency was more important than modulation and coding and that it will be particularly important to avoid modulation and coding schemes which while theoretically spectrally efficient, introduce end to end delay and delay variability.

Modulation and coding also needs to be closely coupled with multiplexing. The three functions are separate but interrelated and together have a fundamental impact on system performance.

No one set of options is perfect. The objective is to arrive at an efficient compromise between spectral efficiency and power efficiency and the optimum choice is likely to be different for the Metre band (300 MHz to 3GHz), Centimetre Band (3 GHZ to 30 GHz) and Millimetre band (30 to 300 GHz).

For example a train travelling at 300 kilometres per hour supported by a radio system at 2.6 GHz will have a constant channel for about one millisecond, equivalent to a sub frame in LTE.

To have an equivalent radio channel at 60 GHz, the train would need to be travelling at 15 kilometres per hour – slower than an elite Marathon runner.

It is therefore clear that adaptive modulation and coding schemes for 5G, particularly 5G mobile wide area systems implemented in the Centimetre and Millimetre bands will need to be different from existing options. If existing options prove capable of being scaled to these higher bands then optimisation will be needed.

Read on

There is significant debate within the vendor community over 5G modulation and coding and multiplexing options and particular options and combinations are claimed to be superior.

The validity of these claims is dependent on the overall performance objectives.

The use of simple GMSK modulation in GSM 2G radio systems for example is not particularly spectrally efficient but performs well in the power domain (efficient power amplifiers). Spectral efficiency is achieved through a combination of bandwidth efficient voice codecs and digital coding gain based on block and convolutional coding.

3 G systems moved to higher order QPSK and 16 QAM modulation. This improves spectral efficiency but performs less well in the power domain. Coding gain is similar to 2G but with more resources (clock cycles of processing) required to equalize the wider (5MHz) channel bandwidth.

4G systems added in OFDM as a mechanism for managing the increased time dispersion of wider bandwidth channels (>5MHz). Time dispersion occurs when a transmitted signal propagates to the receiver via multiple paths with different delays due to the varying path lengths. In the frequency domain, a time dispersive channel translates into a non-constant frequency response. The channel becomes frequency selective. As bandwidth increases, time dispersion increases.

OFDM is useful because it can be implemented using a computationally efficient FFT. Time dispersion is managed by the use of a cyclic prefix in which the last part of the OFDM symbol is copied and inserted at the beginning of the next symbol. This is a time domain process that results in improved frequency diversity.

It also allows for interference coordination for broadcast channels in which identical time aligned signals are transmitted from multiple cells in a single frequency network. At the terminal, inter cell interference appears as time dispersed signal corruption. Provided the cyclic prefix is long enough, broadcast rates are only limited by noise.

Time dispersion can of course be exploited in multi antenna systems. A simple example is two antenna space time coding in which modulation symbols are transmitted on the first antenna then modulated on to the second antenna with the order of the modulation symbols reversed. This is supported within 3G WCDMA systems.

In 4G LTE, the use of OFDM allows for more complex space frequency time coding where a block of frequency domain modulation symbols are mapped to specific frequency sub carriers and then applied to an antenna element.

The performance gain achieved is a function of the relative wavelength spatial separation between the antennas.

If the antennas are relatively close together, for example in Metre band antenna arrays, there will be high mutual antenna correlation (minimal wavelength separation). The transmission beam can be steered by applying different phase shifts to the different antennas but the beam will move slowly and be relatively wide. The result should be some increase in the received signal strength but no additional diversity against radio channel fading.

If the antennas are relatively wide apart in terms of relative wavelength, for example in Centimetre and Millimetre band antenna arrays, there will be low mutual correlation. Combined with polarisation diversity this allows the antenna weights to have complex phase and amplitude values. This enables antenna pre coding in which the phase of the transmitted signal is rotated to compensate for unwanted phase off sets.

The result should be fast beam forming and or highly directive beam forming which together should provide more protection against radio channel fading.

It is also possible to spatially multiplex the antenna array to utilize high signal to noise ratios to support high data rates (MIMO) and or to implement successive interference cancellation.

In successive interference cancellation, the receiver demodulates and decodes one of the spatially multiplexed signals then subtracts that time and frequency signal energy from the next spatially multiplexed signal. The first signal decoded will need to be more robust in order to counter the higher interference floor.

This is usually achieved by using a lower order modulation and lower coding rate. It could also be achieved by applying different power levels to each successive symbol stream. This is usually described in the technical literature as Non Orthogonal Multiple Access (NOMA). The power efficiency and linearity requirements of NOMA require additional study.

The existing LTE physical layer is in many ways elegant. The disadvantage is that the cyclic prefix absorbs bandwidth and power. If the frequency selectivity of the channel exceeds the span of the DFT then the inverse DFT will be unable to reconstruct the original block of transmitted signals. The sub carriers are also sensitive to narrow band interferers. The scheduler will work its way around this but useable bandwidth will decrease.

More problematically the OFDM sub carriers have a habit of ganging up together so large resource block allocations can cause issues with out of band (OOB) emissions. The highest order modulation option at the moment is 256 QAM. The composite waveform, including its envelope variation, is a product of the modulation and multiplexing and coding and scheduling. OOB can be managed proactively by not using resource blocks at the edge of the channel but as with narrow band interference, there will be a loss of useable bandwidth.

There are also doubts as to whether a conventional Fourier transform will be able to scale to the higher frequencies, wider channel bandwidths and higher data rates expected from 5G wide area radio systems.

These limitations provide the basis for alternative modulation and multiplexing schemes including Filter Bank Multi Carrier where each sub carrier is individually shaped in the frequency domain by an individual sub band filter. The claimed benefits are higher out of band attenuation from the sub band filtering with a low post sampling filter rate, better spectral efficiency due to the eradication of the cyclic pre fix and asynchronous allocation of empty sub channels to new users as they become available. This would be particularly useful in dynamic spectrum access schemes. The performance of the sub band filtering will however be crucial to the overall efficiency of this option.

Last but not least there are the various coding options.

The technical literature on coding tends to leap into post graduate maths at the first opportunity but essentially the job of all coding schemes is to answer two questions. Is this a 0 or a 1? And Is this my 0 or 1?

The correct answer to the first question delivers sensitivity gain. The correct answer to the second question delivers selectivity gain.

Seventy years of coding theory can be summarized as follows.

In the beginning, parity checks were used to detect bit errors. If the parity check sum failed the coded bit stream would be retransmitted. Parity checks are still used today.

Block codes developed from parity checks. They work on a similar basis but the parity check number points to where an error has occurred in a code block which can then be error corrected.

Convolutional codes are different because they use memory. A 0 or a 1 travels through the encoder influencing the output code word at every multiplication point. This is analogous to asking is it a 1 or a 0 multiple times. There is no bandwidth expansion. The cost is coding and memory delay and clock cycle overhead as the code stream passes through the convolutional encoder and decoder.

Turbo coders are two convolutional coders working in parallel and use soft decision information, for example instantaneous channel conditions, to weight the decision matrix in the decoder. Each codec encodes the entire input block of data bits.

The patent for turbo codes (held by France Telecom) expired in August 2013.

Low density parity check codes (LDPC) are an alternative to turbo codes and are proposed in several 5 G coding schemes. They use similar decision trellis methods but the data is encoded in short blocks across multiple encoders.

Signal to noise ratio and or signal to interference and noise ratios are measured and described by the Channel State Indicator. As these ratios deteriorate, the coding overhead will increase and lower level modulation options will be used. A 10% error threshold is usually used to determine the choice of modulation and coding.

The objective is to avoid triggering higher level send again protocols. An out of sequence transport block caused by errors or coding error extension will incur an automatic repeat request round trip delay of 8 milliseconds which is tolerable when admission control is being managed at frame level (every 10 milliseconds) but destructive if admission control is being managed at sub frame level (every 1 millisecond).

Modulation and coding requirements for wide area mobility would therefore appear to be different from fixed point to point. By definition there is no need to code fixed point to point for Doppler shift (frequency shift) and there will be minimal time dispersion assuming the link is line of sight.

Combined with the >40 dBi of directional gain available from a dish antenna in the Centimetre or Millimetre bands this allows lightly coded higher order modulation to be implemented.

This however may change in 5G networks. Mobile platforms for example cars, boats and trains and planes could act as repeaters and relays and would need to be channel coded for Doppler shift. Active multiple antenna arrays in the Millimetre band could be producing similar gains to existing dish antennas with existing 1 to 2 kilometre hop lengths being significantly extended. Wide area coverage in the Millimetre band will therefore be based on narrow beam width links with moderate line of sight time dispersion and with near line of sight time dispersion exploited to provide additional delivery bandwidth. The physical layer requirements for wide area mobile coverage and backhaul will therefore become more similar over time.

It is thus not implausible to consider using similar modulation, coding and multiplexing schemes for back haul and wide area mobile connectivity. The implied performance gain and cost reduction potential could be essential to producing an economically sustainable 5G radio system.

About RTT Technology Topics

RTT Technology Topics reflect areas of research that we are presently working on. We aim to introduce new terminology and new ideas to help inform present and future technology, engineering, market and business decisions. The first technology topic (on GPRS design) was produced in August 1998.
http://www.rttonline.com/tt/TT1998_008.pdf

17 years on there are over 200 technology topics **archived on the RTT web site.**

Do pass these Technology Topics and related links on to your colleagues, encourage them to join our **Subscriber List** and respond with comments.

Contact RTT

RTT, the **Jane Zweig Group** and **The Mobile World** are presently working on a number of research and forecasting projects in the mobile broadband, two way radio, satellite and broadcasting industry.

If you would like more information on this work then please contact **geoff@rttonline.com**

00 44 208 744 3163