RTT TECHNOLOGY TOPIC
November 2005

Do we need Super 3G?

A Quarter Century Learning Experience

Over the past 25 years, the cellular industry has had to address four common and inter related challenges - coverage, capacity, consistency (voice quality, call success rate, dropped call rate) and, of course, cost.

It is the role of technology to deliver steady progress in all these areas.

Vendors occasionally promote specific new technologies, smart antennas being one example, which are claimed to deliver step function improvements.

Such claims are always overstated and are generally the direct consequence of marketing getting ahead of engineering reality.

Because of the way in which the industry and the standards making process is structured, there is a step function effect every time a new radio standard is introduced. This effect is however not due to any one specific product but is dependent on the combination of multiple device and design innovations. These step functions usually involve an initial small step backwards in order to achieve a (hopefully not over delayed) large step forward.

It's a sort of traditional technology two step tango.

The introduction of GSM in 1992 initially resulted in an increase in handset and base station component count (a cost penalty) and a decrease in radio system sensitivity (a cost and performance penalty).

It took three years (1995) before GSM phones became generally smaller and lighter and smarter than the analogue phones they were meant to replace. From there on, GSM delivered user, cost and performance benefits that drove the industry through an unprecedented period of global expansion.

Ten years later, the introduction of 3G phones and 3G networks prompted another sashay on to the dance floor for a second attempt at the two step technology tango.

It has to be said that over a period of ten years, the industry's dancing skills had made little discernible progress.

History suggests the same process will be repeated when or if a 4G or 3.5G (Super 3G) transition is attempted.

So perhaps we should examine how the industry's dancing skills could be improved.

How to dance the tango

Dancing depends on having a refined sense of timing, an ability to anticipate and the ability to work with your partner (or partners if we were to use line dancing as an analogy).

A sense of timing translates into not introducing new technology before it works. Easy to say, hard to do as marketing and engineering have, and always will have, communication issues. Taking dance classes is one thing; changing human nature is rather more ambitious.

The ability to anticipate actually implies an ability to size a challenge before it arrives.

In this area, it should be relatively easy to make an improvement.

For example, how to answer the question 'How much offered traffic will there be in 5 years, ten years, 15 years, 25 years, what will be the traffic mix, what will be the impact on technology requirements and (related) business models?'

The trick here is not to rely on subjective points of view but to find a tangible objective basis for projecting future technology and market needs.

The answer revolves around how many people will own phones, how often will they use them, and what will they use them for.

Interestingly, there are a number of non-linear functions involved here.

Offered traffic has traditionally been dimensioned using voice Erlangs and (more recently) a mix of voice and data Erlangs.

However there is anecdotal (and increasingly actual) evidence that as 'session width' increases, 'session length' increases.

'Session width' is just a way of describing the complexity of a session. For instance is it voice, voice and text, imaging and voice, imaging and voice and text, video and voice, video and voice and text, user to user, user to multi user or multi user to multi user.

An increase in dimension along any of these 'application vectors' results in an increase in 'immersion'. As user immersion increases, session length increases.

Telecoms as the new tobacco industry

To divert from our dance floor analogy for the moment, this is telecoms taking over the traditional role of the tobacco industry with a business model based on addiction and dependency and it is this which will drive future offered traffic (and future traffic value).

The addictive parts of the traffic mix will (probably be) imaging bandwidth based and we base this on an analysis of the presently changing form factor and functionality of cellular phones.

For the purposes of this short article, let's oversimplify phone classification into standard phones, smart phones and a new generation of 'superphones' or, perhaps, 'future phones'.

Standard Phones, Smart Phones and Super Phones (FuturePhones).

Standard phones are voice dominant and/or voice/text dominant. Standard phones change the way we relate to one another. These phones have relatively basic imaging and audio capabilities.

Smart phones change the way we organise our work and social lives. They have more advanced imaging and audio capabilities than standard phones.

'Superphones'/Future Phones' change the way in which we relate to the physical world around us.

Superphones/future phones have (or will have) advanced imaging capabilities - for example, >10 Megapixel still imaging resolution, stereo image processing and 3D resolution capabilities.

Note that most humans have two ears **AND** two eyes. We hear and see in stereo and process these inputs to build a 3D map of the physical world around us. Superphones/future phones will do the same.

Superphones/future phones will substantially increase user immersion, which in turn will increase session length, which in turn will increase session value provided that sufficient capacity can be made available at sufficiently low cost at sufficiently low power.

A percentage of this offered traffic will be non real time.

A **surprising percentage** of this offered traffic will be **real time** (the addiction and dependency model depends on instantaneous gratification).

A percentage of this traffic will be downlink biased.

A **surprising percentage** of this traffic will be **uplink biased**. The addiction and dependency model is linked to the other fundamental human driver, vanity and ego. Vanity and ego together will drive user generated content - the uplink is going to be the major contributor to future network value.

Now this might seem a trifle apocalyptic and there are strong counter arguments that the downlink model originally presented and promoted by the UMTS Forum will prevail but it is possible and plausible that downlink asymmetry has been substantially over assumed.

The implications of uplink asymmetry

This has profound implications both for handset and network design.

For a start, any encode process, for example, image coding or video coding in a handset typically uses at least twice as much power as a decoder (it's much harder computationally to compress than to decompress).

In addition, most present wide radio systems are being provisioned on the basis of what might be argued to be over extreme downlink asymmetry, HSDPA being a present example.

First and second-generation radio systems have traditionally been designed with a balanced link budget.

Going to an asymmetric downlink model implies a shift in the wrong direction.

To use our dancing analogy, the dancing partners have headed off to the wrong end of the dance floor.

So it is likely that there will be a shift of focus towards technologies and techniques that can reduce the power drain of the encoders in the handset and a parallel shift towards technologies and techniques that can deliver improved sensitivity and selectivity in the base station (the point of entry into the wireless network).

There are no single killer solutions in this space, more a combination of techniques.

In the phone, attention is shifting to developing algorithmically efficient techniques for compressing multi Megapixel images without loss of perceptual quality together with parallel techniques to improve video capture capabilities.

In the base station, the solutions are partly hardware based (more sensitive radio front ends), partly processor based (active interference cancellation) and partly system based.

The system-based solutions are typically managing admission control algorithms and load balancing both of which help to reduce base station receiver noise and interference.

Spectrum - it's never there when you need it

Spectral efficiency is counter intuitively not as big an issue as the industry generally thinks. There is arguably a glut rather than shortage of radio access spectrum. It's just not always in the right place at the right time.

Again, one of the potentially most effective solutions to these localised issues of a lack of spectrum is to optimise load balancing across wide area, local area and personal area radio systems.

This is the so-called 'seamless convergence' story and makes complete sense from a

technical perspective. At any given time, there will always be a 'broadband best connect' option where there is a single best choice of available bandwidth based on both availability and power efficiency. The choices could include Bluetooth or UWB, WiFi or wide area. Note it is not necessarily an either/or choice and it is very reasonable to look at task allocation across multiple radio access schemes supported simultaneously.

Incidentally, this is a kind of alternative interpretation of MIMO (multiple input multiple output) systems. MIMO systems are traditionally predicated on the idea of exploiting space diversity in an end to end channel using one specific radio standard. It could be argued that extending MIMO algorithms across multiple access schemes might be more effective. Standards diversity rather than space diversity.

Anyway, the point about 'seamless convergence' is that though it may be relatively straightforward technically, the commercial implementation of these inter RAT (radio access technology) handover algorithms is depressingly complex.

For instance, how do you define 'best connect broadband'? Is it 'best connect' from the user's perspective? (Lowest tariff/fastest access/least power consumed) or 'best connect' from the operator's perspective (maximum revenue).

The two are mutually incompatible and it can be more or less safely assumed that this inherent incompatibility will more or less guarantee the failure of present 'seamless convergence' standardisation initiatives.

The eternal quest for power efficiency

So we must look for power efficiency in other areas of which the most promising is the prosaic but fundamental area of adaptive algorithmic optimisation.

This starts in the source encoder where the choice of encoding algorithms is determined by the image statistics (or in voice and audio the audio characteristics of the input waveform).

Philosophically this is the basis of all forward looking source adaptive voice, audio and image codec standardisation. The principles of adaptive coding are writ large for example in the JPEG 2000 and MPEG Part 10 wavelet based compression standards.

The processor or typically processors in a multi processor handset mostly already have task-adaptive frequency and voltage scaling.

At the physical layer and MAC layer, almost all forward looking technical standards are adaptive in the sense that either the data rate, the modulation, the channel coding and power allocation can respond either to changes in the radio channel and/or instantaneous or near instantaneous changes in the bandwidth requirements of the supported application.

The addition of an OFDM multiplex in Release 7 (HSOPA) and the use of OFDM in future 1XEVDO radio systems (Media Flo) extends this principle by introducing

scalable bandwidth.

As an example, if you take all present and proposed OFDM radio systems, they scale from bandwidths of 9 kHz (Digital Radio Mondiale) to 528 MHz (multi band OFDM UWB). User bandwidth scalability is anything between a few bits per second (low power telemetry) to 480 Mbps (wireless USB), modulation schemes scale from 2 level FSK to 64 level QAM.

Other areas where adaptive schemes yield performance gain include displays. Backlights that respond to ambient light level are now common place but more aggressive adaption schemes addressing refresh rates will likely yield significant future power efficiency gains. Bi stable displays that use no power at all if the information remains unchanged are one example.

Who needs a Super 3G Solution?

Given all this, you might ask why we need Super3G and the answer is we don't. We already have the solutions we just need to apply them.

And what is the biggest problem that we need to solve?

It's power drain. Power drain constrains offered traffic bandwidth and offered traffic value.

Probably the most significant power savings are likely to be realised by an area of emerging research work now often described as cross layer coding. This is where instead of looking at one problem area in isolation; the coding schemes from the application layer down are cross optimised. Source coding is matched to source statistics expressed in meta data descriptions that are matched to optimised IP traffic shaping protocols that are matched to optimised PHY and MAC radio access systems. (More on this fascinating subject in future Hot Topics).

Getting to know your partner

Which brings us back to our dancing analogy.

If these cross layer cross disciplined schemes are to stand any chance of success, then engineering teams from different disciplines will need to work together more effectively and marketing/engineering communication will need to be substantially improved,

The technology two step tango will remain as a major part of the future dance repertoire but we can reduce its effect by avoiding the relatively abrupt transitions that the industry went through in the AMPS/ETACS to GSM transition and the present GSM to 3G transition.

Vendor vanity might still lead us down the path of adopting a 'new' super 3G standard, which for argument would have significant differences when compared to present Release 7 proposals. This would be ill advised. The old argument that a new standard puts clear water between Tier One vendors and other smaller competitors

no longer really applies given that much of the potential performance differentiation has now moved into the system solution domain (a combination of small but significant optimisation techniques).

In this context, evolution is better than revolution (and certainly safer).

Failing this, just remember it takes two to tango, so choose your partners wisely and well.

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 clarify present and future technology and business issues.

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

Contact RTT

RTT, the Shosteck Group and The Mobile World are presently working on a number of research and forecasting projects in the cellular, 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