RTT TECHNOLOGY TOPIC
July 2006

HSPA Device and System Efficiency

Introducing this month's Hot Topic

In this month's Hot Topic we explore how HSDPA devices and HSDPA system protocols can potentially increase radio system efficiency, how radio system efficiency may translate into cost saving and revenue gain and how cost saving and revenue gain may translate into an increase in spectral value.

We use HSDPA as an example. Actually we use HSDPA (the downlink) and HSUPA (the uplink) which together should be **described generically as HSPA** - high speed packet access. However the metrics apply to any wide area mobile wireless adaptive radio system.

In the process we set out to illustrate the benefits of a closer integration of HSPA technology, engineering and market and business planning.

Defining HSPA Spectral value

Spectral value is a composite of the cost of the spectrum, the cost of infrastructure to access the spectrum, the cost of user devices and the revenues generated from those devices. The performance of the infrastructure in terms of radio and network efficiency and the performance of user devices in terms of their bit rate, range and power efficiency will have a direct influence on revenue. The process is essentially the translation of silicon value into system value into spectral value. A valid business model requires this value to be greater than the amortised cost of the spectrum.

HSPA spectral supply and its impact on HSPA cost and revenue assumptions

Spectrum may be allocated or auctioned at a rate that is slower or faster than the market can absorb, creating a spectral shortage or a spectral glut. The balance of spectral supply and demand therefore influences the cost saving and revenue gain assumptions that operators and their vendors use to justify HSPA technology and engineering investment.

Figure 1 Present and potential future HSPA cellular spectrum allocations from 800 MHz to 2.6 GHz

Band	3GPP	Allocation	Uplink	Downlink	Region
I	2100	2x60 MHz	1920-1980	2110-2170	Present UMTS
II	1900	2x60 MHz	1850-1910	1930-1990	US PCS
111	1800	2x75 MHz	1710-1785	1805-1880	GSM Europe, Asia, Brazil
IV	1700/2100	2x45 MHz	1710-1755	2110-2155	New US
V	850	2x25 MHz	824-849	869-894	US and Asia

VI	800	2X10 MHz	830-840	875-885	Japan
VII	2600	2x70 MHz	2500-2570	2620-2690	New
VIII	900	2X35 MHz	880-915	925-960	Europe and Asia
IX	1700	2x35 MHZ	1750-1785	1845-1880	Japan

As Figure 1 shows, there are already nine band allocations or proposed allocations for cellular spectrum between 800 MHz and 2.6 GHZ. These are specifically bands that are either already designated for UMTS 5 MHz channels or could be used for UMTS and are therefore by definition, suitable for use with HSPA.

The table excludes the 450 MHz band and 700 MHz band. The 450 MHz band is arguably not suited to 5 MHz channel spacing. The 700 MHz band could be but remains subject to further discussion from a regulatory perspective. Either way, this suggests that the potential exists to produce a **Deci Band HSPA handset** capable of accessing over 800 MHz of cellular spectrum.

This illustrates a point that we often make, that the cellular industry is not intrinsically short of spectrum but is short of power and quite often short of money. Power and money are interrelated and can be expressed in engineering terms as an HSPA radio link budget.

The radio link budget as a measure of system efficiency and cost
Engineers use link budgets as the basis for calculating network density (range or coverage) and spectral efficiency (capacity). The radio link budget is a composite of many factors each of which has an impact on system efficiency and therefore on cost. For example;

Lower frequencies deliver better range.
The radio link budget at 800 and 900 MHz will be at least 8 or 9 dB better than the radio link budget at 2 GHz so coverage will be substantially better for a given network density.

Higher frequencies deliver more capacity
However the higher the frequency, the better the frequency re use so the higher bands will yield more capacity per MHz of spectrum.

Smaller guard bands decrease sensitivity
The guard band is the amount of distance between the top of the mobile transmit band and the bottom of the mobile receive band. From Table 1 it can be seen that the amount of guard band between transmit and receive frequencies varies from 20 MHz to 355 MHz. Smaller guard bands translate into a decrease in receive sensitivity.

Smaller duplex spacing decreases sensitivity
A different but similar issue. From Table 1 it can be seen that duplex spacing varies in the examples shown between 45 MHz and 400 MHz. Less duplex spacing translates into a decrease in receive sensitivity. However:

Extended duplex spacing increases sensitivity but complicates and potentially compromises channel measurement and channel feed back processes

A larger than normal duplex spacing, for example, the 400 MHz duplex spacing used in US AWS Band IV, implies a large difference in propagation behaviour between the uplink and downlink which makes it harder to implement accurate uplink and downlink channel measurements . Accurate channel measurements are needed in order to support adaptive modulation and coding schemes and admission control algorithms. If these are not optimised, system efficiency will suffer.

The practical impact of different guard bands and non standard duplex spacing
The 3GPP specifications allow a 2 to 3 dB relaxation in sensitivity depending on the guard band and duplexing used. **Every 1 dB of sensitivity lost translates into a 10 % increase in network density.** This is a significant metric that needs to be factored into capital and running cost calculations.

An HSPA link budget is different
An HSPA link budget is different from a conventional link budget in that it is heavily influenced by system implementation, the level of service provided to customers and the traffic mix, described below as 'content'. Some of these system gains (and occasionally system losses) are hard to quantify, as we shall now explain.

HSPA capital and running cost calculations
Figure 2 Coverage and capacity and the impact of customers and content on cost calculations

Coverage	Capacity	Cost	Customers	Content
		Revenue gain		

We have shown that frequency allocation has a direct impact on coverage and capacity. Coverage and capacity are part of a cost calculation. Both have a direct impact on potential revenue. A user out of radio range is out of billing range, a blocked call is (often) a lost call

Costs and margins can be expressed simplistically in terms of dollar per byte delivered but cost calculations are sensitive to the assumptions used for user service level, traffic loading and traffic distribution.

For example, an operator might choose to dimension network capacity to support a busy hour loading of 80%. This might however violate acceptable user service levels. **A violation of user service level could** either surface as a blocked or dropped call (conversational voice and video), an over length buffer or re buffering in a streamed session, unacceptable latency and delay variability in an interactive session and/or unacceptable packet delay or packet loss in a best effort session. Dropping the busy hour loading to 40% will resolve these service level issues but will **double the delivery cost.**

Coverage and delivery cost are therefore determined by individual customer service levels and the content and traffic mix delivered to and from the user.

Margin is obviously a function of cost per delivered megabyte but also a function of any revenue gain achievable from additional services.

This also illustrates the point that although we might not be intrinsically short of bandwidth, it is not always in the right place at the right time, hence the need for device and system efficiency gain. It may also need to be configured in a way that allows us to develop new user service propositions. An example would be the use of multiple code channels on the uplink and downlink to deliver multiple content streams. The additional complexities implicit in these channel configurations will have an impact on system efficiency.

Device Performance and the HSPA Link Budget

The radio link budget for HSPA is a composite of the sensitivity, selectivity and stability of the handsets and base stations, the propagation loss and propagation properties of the radio channel and a whole range of system parameters that can be changed in response to changing channel and operational conditions. This process of adaptation at system level is the mechanism by which HSPA and other packet based radio systems deliver gains in system efficiency and power efficiency.

Arbitrarily we suggest that sensitivity, selectivity and stability are determined by 'technology factors', for example, device performance. We have said that the band allocation (guard band and duplexing) makes a significant difference. In addition, other system factors have a major impact on how the link budget is distributed. For example, reducing the load on a base station will lower the noise floor of the base station which will increase the uplink range.

Let's start by working through some of the device specific technology variables in the table below

Figure 3 Sensitivity, selectivity and stability.

Sensitivity	Selectivity	Stability
Handset receive sensitivity	Handset receive selectivity	Handset frequency, phase and amplitude stability.
Base station receive sensitivity	Base station receive selectivity	Base station frequency, phase and amplitude stability
Advanced handset receivers	Advanced handset receivers capable of working with advanced BTS/node B transmitters	
Release 5 Single rake receivers		
Release 6 Enhanced Type 1 Rake receiver with receive diversity	Release 6 Enhanced Type 2 Single Equaliser receiver	
	Release 7 Enhanced Type 3 Equaliser receiver with receive diversity	

Handset sensitivity
This will vary from band to band and from handset to handset. We have said that there is a 2 or 3 dB difference depending on the band being used. There will be a similar difference between different handsets from different manufacturers. The mix of handsets being used on a network at any given time therefore has a fundamental impact on network performance.

The Release Standard to which the handsets comply will also make a difference. This introduces us to the topic of **Advanced Receivers**.

Advanced Receivers
Initially HSPA handsets are likely to be **3GPP Release 5** compliant which means that they are **single rake receivers.**

Release 6 introduces **Enhanced Type 1 receivers with receive diversity.** This involves adding a second antenna to the handset and implementing a dual receiver front end to support a combining function. A diversity receiver is based on the principle that only one of the antennas will be in a deep fade at any given moment. In practice, it is hard to get sufficient spatial separation in small form factor handsets particularly in the lower frequency bands so the real benefits may be marginal in practice.

Base station sensitivity
The same story applies though as we have said above, the noise rise of the base station can be managed at system level.

Handset selectivity
This is the ability of a handset to separate wanted signal energy from unwanted signal energy. The handset needs to achieve user to user selectivity and code to code selectivity.

Code to code selectivity may involve the de correlation of a single code or simultaneous de correlation of a 5 code, 10 code or 15 code multiplex.

Each code represents a phase argument. Given that the codes are being delivered over a multi path channel, this phase argument will have been compromised, a process described as a loss of orthogonality.

Advanced Receivers that work with Advanced Transmitters
Enhanced Type 2 and Type 3 receivers are receivers optimised for use with advanced transmitters. Advanced transmitters are base stations that are supporting users with multi code (5, 10 or 15 code) downlinks. The process of equalisation restores the orthogonality of the codes.

An **enhanced type 2** handset receiver is a **single equaliser receiver with diversity** gain. This receiver will deliver a benefit in terms of multi code de correlation performance, in other words, a gain in **selectivity.**

An **enhanced Type 3** handset will be an **equaliser receiver with receive**

diversity. This should deliver a performance gain both in terms of **selectivity** (the ability to de correlate multi codes delivered over a time variant channel) and **sensitivity (**energy combining as a product of spatial diversity).

Enhanced Type 2 and enhanced type 3 receivers will both **depend on accurate real time channel estimation** and discussions are still ongoing as to how the control channels will be configured. The theoretical gains achievable from these advanced receivers will therefore be dependent on well implemented system level closed loop feed back control which may be particularly hard to implement in networks with non standard duplex spacing.

Base station selectivity
The base station has a harder job to do in terms of selectivity in that it could be seeing possibly hundreds of mobiles each delivering single or multi code uplink transmissions. Future base station platforms will implement various forms of **multi user detection** and **multi user interference cancellation** which should deliver incremental gains in uplink selectivity.

Handset Stability
This is the ability of the handset to perform receive and transmit functions over a wide range of dynamically changing channel and operational conditions. For example modulation accuracy, described as error vector magnitude (EVM) is an issue in the transmit chain particularly with multi code uplink transmission. EVM is also partly a function of the linearity of the transmit chain. Delivering sufficient linearity to support multi code transmission efficiently and accurately remains a significant device and design challenge.

Base station Stability
This is the ability of the base station to perform receive and transmit functions over a wide range of dynamically changing channel and operational conditions. For example modulation accuracy, described as error vector magnitude is an issue in the transmit chain particularly with multi code downlink transmission.

Advanced Base Station Transmitters with Multiple Transmit Antennas
Additionally, **Release 7** may include specific recommendations on transmit spatial diversity, also known as **MIMO (multiple input/multiple output) systems** where multiple code channels are sent on separate antennas. The real throughput benefits of MIMO remain to be proven particularly in larger cell deployments. The processing load in the handset will also be substantial.

System Performance and the HSPA Link Budget
In practice, it may be easier to deliver throughput and power efficiency gains by optimising system level performance.

Working through some of the system specific engineering variables in the table below

Figure 4 Adaptive processes used in HSPA (system level engineering variables)

Adaptive frequency	Inter frequency handover, inter band handover, shared/dedicated channels, and inter system handover.

Adaptive modulation	QPSK/16 QAM
Adaptive channel coding	1/7 to 1/1 coding
Adaptive power	10-15 dB downlink, 71 dB uplink
Variable bit rate	Bits per TTI (transmission time interval)
HARQ retransmissions	Chase combining/incremental redundancy
Microbuffering	Conversational content
Macro buffering	Streamed, interactive, best effort
Micro scheduling	.5, 2 or 10 ms TTI
Macro scheduling	Busy hour traffic management
CQI context	Scheduling priority indicators, discard timers and CQI

Adaptive frequency
This involves the integration of existing handover algorithms with HSPA specific algorithms. Early implementations of packet radio, for example GPRS, attempted to get away with 'break before make' handovers. This proved non optimum for most wide area mobile applications and prompted a general move back to network assisted handover in which any momentary discontinuity in terms of throughput is made invisible to higher layers in the protocol stack. An example would be the implementation of an algorithm in the RNC which predicts when a handover is going to take place, managing the packet flow between the serving and handover candidate base stations to either avoid or certainly minimise packet loss or unacceptable packet delay.

Shared/Dedicated Channels
Early implementations of HSPA are generally predicated on a 5 MHz channel being shared between Release 99 channels used for conversational voice and video and HSPA channels. This is not efficient. Sharing bandwidth implies sharing power and managing this sharing process in an environment in which both the radio channel and traffic loading are highly dynamic. This implies a signalling load which is generally uneconomic both in terms of spectral and power efficiency. As HSPA handset data rates increase it will be far more efficient to implement a dedicated HSPA data channel.

Inter system handovers
Inter system handovers include HSPA to Rel 99 handover, handover to and from GSM and potentially handover to and from other available 'best connect' radio bearers, for example WiFi. Inter system handover optimisation can have a major impact both on overall spectral efficiency and overall power efficiency (handset duty cycle).

Adaptive Modulation
The principle of adaptive modulation is that **downlink** data rates can be increased in favourable channel conditions.

Adaptive Channel Coding
The principle of adaptive coding is that uplink and downlink user data rates can be

increased in favourable channel conditions. An effective code rate of 1/7 means that 137 user bits are carried using 960 channel bits per transmission time interval per high speed packet data shared channel. An effective code rate of 1/1 means that 960 user bits are carried per 960 bits sent.

Adaptive Power Control
Rel 99 uses fast power control in the uplink and downlink. HSPA uses fast power control only in the uplink and manages slow and fast fading on the downlink by micro scheduling (see below).

Variable bit rate
Both the hand set and the base station can change the number of bits sent each transmission time interval.

HARQ retransmissions
These are described as **Fast Layer 1 retransmissions** and are managed between the handset and the base station. Retransmissions are where a parity check has failed and the receiver asks for the bit burst to be re transmitted. Either the same bits are sent multiple times (**chase combining**) or different sets of the original bit burst are sent multiple times (**incremental redundancy**). These processes have the advantage of being self adaptive and do not depend on the feed back control required for each of the other adaptive processes listed above and below. Some of the technical literature suggests that HARQ increases range and in building penetration.

Microbuffering
This is buffering that is sufficiently fast to be invisible to higher layers of the protocol stack. An example would be the buffering of one or more short (2 millisecond or .5 millisecond) transmit frames within a 10 millisecond frame.

Macrobuffering
This includes the buffering of multiple 10 millisecond frames and may extend to several seconds depending on the buffer bandwidth available in the handset, base station or RNC. Macro buffering can be used to smooth offered traffic either across the radio channel or across the IUB and IU interfaces.

Micro scheduling
The use of 2 millisecond or .5 millisecond frame lengths mean that potentially the base station can schedule downlink traffic 'on top of the fades' at least for slow mobility users who are relatively close to the base station. Short duration 2 millisecond or .5 millisecond frames become very expensive in terms of signalling overhead for users closer to the edge of the cell but show efficiency benefits for high data rate users close to the base station.

Macro scheduling
Macro scheduling can be spatial or temporal or both. Spatial macro scheduling is where the admission control algorithm gives users who are close to the base station preferential access to channel bandwidth. This maximises throughput efficiency in the cell but can end up with users at the cell boundary never getting any throughput. The answer is to use a technique known as proportional fair scheduling which balances cell throughput optimisation with user throughput requirements designated as a

scheduling priority indication. Temporal macro scheduling is where a network shifts best effort traffic out of the busy hour.

CQI Context

All of these adaptive processes with the exception of chase combining and incremental redundancy are dependent on getting an accurate context upon which to base admission control and scheduling decisions. This is described in HSDPA as the Channel Quality Indicator (CQI) and is based on an index of 31 levels or steps each equating approximately to a 1 dB step in the high speed data shared channel signal to noise and interference ratio. (SINR).

However the admission control algorithms have to take this information and qualify an admission decision on the basis of a number of other factors including the user specific channel specific scheduling priority indication.

Although it is possible to change any of the adaptive parameters just listed every Transmission Time Interval, which means theoretically every 10, 2 or .5 milliseconds, it takes typically 6 milliseconds to respond to a change in CQI so in practice these adaptive processes are implemented every 10 millisecond frame. This means that HSPA as presently implemented is to all intents and purposes an ATM radio layer though it is never described as such in the technical literature.

The implications of HSPA device and system performance for HSPA market and business planning

As with most adaptive radio systems, HSPA is capable of delivering significant spectral and power efficiency benefits but the realisation of these benefits will be dependent on the integration of technology, engineering, market and business planning.

For example, many of the system gain mechanisms described above are based on the differential scheduling of users and/or the differential scheduling of user specific content streams. This has profound implications for user experience management, customer relationship management and billing.

If service level agreements become over complex, many of the cost savings achieved will disappear and re appear in the customer support domain.

Operators that recognise and successfully manage this subtle but significant coupling between HSPA technology, engineering, market and business management will potentially achieve a competitive market and business advantage.

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