

UNCLASSIFIED

QinetiQ

This document has been produced by QinetiQ, Defence and Technology Systems for Ofcom under contract number 410000262 and provides an evaluation of software defined radio.

An Evaluation of Software Defined Radio – Main Document

Editor: Dr. Taj A. Sturman
QinetiQ/D&TS/COM/PUB0603670/Version 1.0
15th Mar 2006

Requests for wider use or release must be sought from:

QinetiQ Ltd
Cody Technology Park
Farnborough
Hampshire
GU14 0LX

Administration page

Customer Information

Customer reference number	N/A
Project title	An Evaluation of Software Defined Radio – Main Document
Customer Organisation	The Office of Communications (Ofcom)
Customer contact	Ahmad Atefi
Contract number	410000262
Milestone number	Of/Qi/002
Date due	March 2006

Editor

Taj A. Sturman	MAL (801) 5378
PB315, QinetiQ, St. Andrews Rd, WR14 3PS	Tasturman@qinetiq.com

Principal authors

Alister Burr	University of York
Julie Fitzpatrick	QinetiQ
Tim James	Multiple Access Communications Ltd.
Markus Rupp	Technical University of Vienna
Stephan Weiss	University of Southampton

Release Authority

Name	Ian Cox
Post	Business Group Manager
Date of issue	March 2006

Record of changes

Issue	Date	Detail of Changes
Version 0.1	05 th Aug 2005	Creation of initial document including structure.
Version 0.2	25 th Aug 2005	First draft for review.
Version 0.3	5 th Mar 2006	Incorporation of reviewed comments.
Version 1.0	15 th Mar 2006	First Issue.

Executive Summary

This document provides an evaluation of software defined radio (SDR). In an SDR, some or all of the signal path and baseband processing is implemented by software, normally in the digital domain, that is, the term SDR refers to *how* the lower layer functionality is implemented. Crucially, the low level functionality of the radio in an SDR can be altered through changes to the software, without any *physical* changes to the hardware. The realisation of SDRs represents a significant step in the evolution of radio.

Two documents have arisen out of this study: An Overview Document and the Main Document (which is this document). It is the intention of this Main Document to provide a relatively detailed account of a particular subject and corresponding references for access to additional information on a particular subject. Due to the sheer volume of this Main Document, each chapter is self-contained and therefore, whilst this gives rise to some repetition, it should enable some ease in dealing with a particular subject of interest.

Considering the processes involved in the reception (or the inverse processes for transmission) of information from the point of reception of electro-magnetic waves in the radio channel to the corresponding information for the user, the following system-related issues are involved. An antenna system is required to handle the air/electrical interface. This might range from a single, passive antenna to an array of 'active', i.e., reconfigurable, antennas controlled by software. The interface between the analogue and digital domains will be handled by analogue-to-digital converters (ADCs) on the receive side and digital-to-analogue converters (DACs) on the transmit side.

Due to limitations in these devices and for the foreseeable future at least, the radio frequency (RF) front-end, incorporating RF amplification and limited frequency translation, will still have to be implemented in the analogue domain (albeit under the control of software). Once in the digital domain, an SDR might use a device such as a field-programmable gate array (FPGA) to implement certain 'high-speed' functions, such as mixing, filtering and sample rate conversion and/or 'general purpose' processors or digital signal processors (DSPs) to implement 'low speed' functions.

The demodulated data are passed to/received from higher level 'application' software, as in more conventional radio architectures. Note that, although the higher level functionality might be implemented in software, this does not automatically classify a radio as an SDR; for a radio to qualify as an SDR, it is the low level, i.e., physical layer, data link layer and parts of the network layer functionality that must be implemented in software.

UNCLASSIFIED

In its simplest form, an SDR might emulate the function of a conventional radio transceiver, replacing some or all of the analogue frequency translation, filtering, modulation and/or demodulation functions with software defined equivalents. An example of this might be the implementation of a stereo frequency modulation (FM) receiver using no more than a high-speed ADC, an FPGA and an audio DAC. A longer term vision of SDR is the development of generic radio platforms that can be reconfigured ‘on-the-fly’, possibly by means of over-the-air downloads, to target multiple radio standards operating over a wide range of carrier frequencies. Realisation of this long-term vision might ultimately lead to the emergence of cognitive radios (CRs), radios that intelligently combine an awareness of the local radio environment with the requirements of the user to reconfigure themselves dynamically to provide the most appropriate and cost-effective communications link possible.

Note that a common misconception is to interchange the terms SDR and CR. Put into a layperson’s terms, the term ‘SDR’ simply refers to a flexible, reconfigurable radio platform whose operation is defined in software. In itself, an SDR has no intelligence. The term ‘CR’ refers to a high-level control entity that intelligently manages the configuration and operation of the radio. It is not a requirement that a CR be implemented using an SDR, or that an SDR has cognitive capabilities. Having said this, SDR is likely to be an enabling technology for CR because of the potential that SDR provides for flexible, reconfigurable radio platforms to be realised.

It is further noted that it is not a requirement of an SDR that it be reprogrammable after manufacture, or even that it must support multiple radio standards and/or true wideband operation. In the case of the latter, as stated above, true wideband operation will be dependent on more conventional, analogue architectures, for the foreseeable future anyway. That is, most SDRs effectively will consist of a flexible, analogue RF front-end, with digital domain processing interfaced at some intermediate frequency (IF).

Ofcom, the United Kingdom’s (UK’s) communications industry regulator, has acknowledged the emergence of SDR and is keen to gain a better understanding of the technologies behind SDR and the possible ramifications for the regulation and management of spectrum usage in the future. To this end Ofcom has commissioned a study to consider key aspects of the SDR technology. The study was awarded to a consortium of five members. Headed by QinetiQ, this consortium also includes Multiple Access Communications (MAC) Ltd, the University of Southampton, the University of York and the Vienna University of Technology.

This document is the culmination of three phases of study, which in broad terms, addresses the following three key groups:

- Technology Requirement Enablers
- Assessment of SDR
- Spectrum efficiency gains of SDR and CR.

In the context of subject categories, these three key groups fall under the following headings:

- Antennas
- RF (Radio Frequency) Linearisation
- Antenna Processing

UNCLASSIFIED

- MIMO (Multiple-Input, Multiple-Output) Technology and SDR
- Waveforms
- Software Aspects
- Security
- Radio Management and CR
- Regulatory Issues
- Commercial Drivers
- SDR: An Assessment of First Applications and Areas of Deployment
- Spectrum Efficiency Gains of SDR and CR.

Of these, QinetiQ was responsible for the antenna, software aspects, security, radio management including CR and an assessment of first applications and areas of deployment for SDR; MAC Ltd considered the regulatory issues, commercial drivers and the spectrum efficiency gains of SDR and CR; the University of Southampton investigated antenna processing and waveforms; the University of York provided the discussion on MIMO technology and the Vienna University of Technology considered RF linearisation. Of the topics investigated, we note that some, namely antennas, RF linearisation, MIMO and cognitive radio and radio management, are not necessarily specific to SDR, or even a requirement of SDR. However, these topics become relevant when considering the longer-term vision for SDR. This report represents the culmination of these essential subject areas.

Antennas

Traditional radio systems are limited to transmission and reception in narrow bands within the RF spectrum. Whilst this has advantages in minimising interference to users in neighbouring bands, it restricts the flexibility of individual transceiver units since they tend to be application-specific, e.g., a portable FM radio receiver (operating at ~100 MHz) cannot receive mobile telephony signals (operating at, say, 1900 MHz).

From a general antenna perspective, these narrowband systems relax the constraints on antenna design; it is extremely difficult to design a truly wideband antenna, e.g., an antenna that can operate, say, across the range of 20 MHz to 2000 MHz. It is likely that commercial, ‘wideband’ SDRs, as with current multi-band handsets, initially at least, will be restricted to operating in multiple, distinct frequency bands specified at design time. Looking towards the future (the next ten years, say), antenna solutions will have been identified that might be reconfigured on-the-fly to adjust dynamically the resonant frequency, gain and polarisation of the antenna, thus leading towards the realisation of true wideband antennas for SDR applications.

As SDR systems seek to optimise the utilisation of the RF spectrum, further advantages can be gained from using ‘smart antennas’. Smart antennas are designed to enable the radiating characteristics to be altered dynamically. Smart antennas potentially could increase system capacity by maximising the gain in the direction of the intended target whilst, in certain implementations, simultaneously minimising the gain in the direction of other users. Thus, transmit power potentially can be reduced, as can the interference power received from and caused to other users. Two main smart antenna technologies have been investigated, namely, beam-steering antennas and beam-switched antennas.

UNCLASSIFIED

The beam-steering approach uses phased antenna arrays to create single or multiple beams and to steer them dynamically to track mobile users. Furthermore, the characteristics of each beam may be altered to account for the migration of user numbers in some areas. Phased arrays provide spatial diversity in that the main beam gain can be directed in a chosen direction, whilst simultaneously placing nulls in other directions to suppress interference to and from other users. In military applications, nulls might also be steered to suppress hostile jamming signals. An approach based on solid state plasma technology has been considered.

The beam-switched approach uses an antenna array comprising a number of static, directional beams. As the intended target moves within the coverage of the antenna system, the beams are switched so that only those beams ‘illuminating’ the target are used. This approach is generally simpler than beam-steering systems, both in terms of signal path complexity and the fact that it does not require the dynamic tracking of users. However, beam-switching implementations are typically less flexible and require a handover scheme to facilitate the switching between beams.

When considering the vision of highly flexible, true wideband SDRs, we have concluded that the antenna design currently provides a significant design challenge. However, emerging novel antenna solutions and state-of-the-art modelling suites, which allow antenna design to become an integral part of the system design, promise to realise significant advances in the capabilities of antennas for SDR applications in the future. A table with estimates for antenna development is provided which quantifies the impact of the challenges and forecasts when novel antenna solutions might emerge.

RF Linearisation

The power amplifier is a critical part of any radio transceiver; its function is to boost the power of the information-bearing signal without introducing significant distortion. There are two sources of distortion in a power amplifier:

- Nonlinearities in the amplifier’s transfer function. By its nature, the power amplifier typically is responsible for a significant portion of a radio’s power consumption and, as such, every effort is normally made to make the power amplifier as efficient as possible. Power amplifier efficiency is generally improved by using either nonlinear designs or using linear designs in or near saturation. Employing amplifiers with nonlinear characteristics to amplify linear signals not only reduces signal fidelity, it can also cause significant out-of-band emissions. Pre-distortion techniques aim to modify the signal input to an amplifier such that the output signal is free from nonlinear effects. The use of such techniques allows more efficient power amplifier designs to be used.
- Linear, temporal and frequency dispersive effects. Dispersive effects in power amplifiers are caused by internal memory effects. In narrowband systems dispersive effects typically can be neglected. However, in wideband systems, the effects of dispersion can be significant. Whilst analogue linearisation methods exist (see Section 3.4.1), they are unable to compensate fully for linear dispersion effects. However, digital pre-distortion techniques are able to counteract the effects of both nonlinearity and linear dispersion (see Section 3.4.2).

UNCLASSIFIED

Most modern wireless systems employ linear modulation schemes. Therefore, linear behaviour at the output of the power amplifier in any ‘generic’ SDR platform is essential. Furthermore, as signal bandwidths increase, measures to compensate for linear dispersive effects will also be a requirement. This is especially true in the case of generic platforms when it will not be possible to optimise the RF front-end for any one particular radio system. Luckily, the inherent flexibility of SDRs makes them well suited to the implementation of advanced, adaptive pre-distortion algorithms in the digital domain. Linearisation methods incorporating both static and dynamic amplifier models have been discussed. In addition, simulation results that clearly show the performance to be gained by implementing such techniques have been presented.

It has been concluded that SDRs will enable high-performance, adaptive RF linearisation schemes to be implemented in the digital domain. Such schemes will be able to compensate not only for nonlinear effects, but also linear dispersive effects. The implementation of high performance linearisation schemes will enable more efficient amplifier designs to be used, even for systems employing wideband, linear waveforms. Not only do such techniques have the potential to offer power savings in the transmitter, from the regulator’s point of view improved signal integrity should help minimise out-of-band emissions, which, in turn, should allow the RF spectrum to be used more efficiently.

However, adaptive linearisation algorithms will need to operate continuously in order to adapt to the continually changing characteristics of the power amplifier. Therefore, it will be crucial that the SDR hardware architectures are designed to be as flexible and as efficient as possible in order to enable suitable linearisation algorithms to be implemented with minimal complexity and power consumption requirements. A table has been generated which indicates epochs for the development and the maturity for SDR technologies related to RF linearisation.

Antenna Processing

Current limitations in converter technology mean that most SDR systems will require some analogue circuitry to convert a received analogue RF signal to a more suitable IF signal for sampling. Similarly, on the transmit side, up-conversion and amplification in the analogue domain will still be necessary. We refer to this manipulation of the signal between the antenna and the digital domain as antenna processing. Different candidate schemes have been reviewed.

The conversion between analogue and digital domains on the receive side and digital and analogue domains on the transmit side are particularly vital. Background converter theory and practical implementations have been discussed. A review of state-of-the-art, commercial ADCs and DACs reveal that the bottleneck resides in the ADC stage; state-of-the-art ADCs generally fall behind the latest DACs, both in terms of sampling rates and signal resolution. To compound this issue, we noted that the requirements, both in terms of sampling rate and resolution, are typically greater for the ADC in an SDR.

UNCLASSIFIED

Concentrating on commercial ADCs, the development of devices based on traditional Si CMOS technology has led to a slow increase in performance of typically 1.5 bits, at any given sampling frequency, over a period of six years, as noted by Walden (Walden R H, "Performance Trends for ADCs", *IEEE Communications Magazine*, vol. 37, no. 2, pp. 96-101, Feb. 1999). In addition to this increase, emerging technologies and techniques, such as the time interleaving of several ADCs, SiGe based circuits, optical sampling and fast logic based on superconducting devices have been discussed.

We might expect such devices to help further shift the performance bounds of commercially available ADCs in the future. However, we note that these high-performance devices are likely to be restricted to base station applications, initially at least, due to power consumption considerations.

Several tables have been provided indicating expected epochs for the development of the technologies associated with antenna processing. Included within this discussion are the epoch expectations for the sampling rate for ADC with 6 - 8 bit resolution and CMOS power consumption epoch expectations, which tend to suggest a future need for new technologies for such devices.

MIMO Technology

By employing multiple antenna elements at both the transmitter and receiver, MIMO technology has the potential to greatly enhance capacity in fading channels by exploiting the effects of multipath propagation. In particular, MIMO potentially provides improvements in diversity due to independent fading of paths between different elements, and improvements in capacity by simultaneously sending different data over different spatial paths or 'pipes'. In principle, capacity might be increased by up to an order of magnitude. Therefore, it is likely that future wireless standards will incorporate MIMO techniques, and hence, it is essential that SDR implementations allow for this.

MIMO is not a requirement of SDR. Similarly, SDR is not essential to the implementation of MIMO; MIMO is simply another 'function' that might be implemented using SDRs (albeit a potentially important one). The potential flexibility of SDR should be well suited to the implementation of MIMO; consequently, SDR ultimately might become an enabling technology for MIMO. Indeed, MIMO and SDR might be considered complementary technologies.

The introduction of MIMO into wireless standards has been considered. Currently, the only major wireless standard to incorporate MIMO is space-time transmit diversity (STTD) and time-switched transmit diversity (TSTD) in UMTS Release 99. MIMO is being considered in 3GPP mainly for the HSDPA extension of UMTS. Finally, MIMO is certain to be included in the new high-speed (>140 Mbit/s) wireless local area network (WLAN) standard being developed by IEEE 802.11 TGn, 802.11n.

There are two main aspects to be considered in the implementation of MIMO in SDRs, namely the RF hardware (including the antennas and the RF/IF/conversion chains) and baseband processing. In addition, discussed within this section is the issue to what extent hardware might need to be duplicated per antenna as well as the impact on complexity at baseband due to adding additional transmit/receive signal paths.

UNCLASSIFIED

In principle, a completely separate RF chain (including low-noise amplifier (LNA)/high-power amplifier (HPA), mixers, IF amplifiers and ADC/DAC) is required for each antenna element. Although some processing might be possible at RF using analogue techniques, such an approach invariably would lead to a suboptimal implementation. One possible approach to optimising the hardware might be to multiplex multiple signals on one, higher bandwidth chain. However, savings in component count might be tempered by the increased requirements resulting from the increased signal path bandwidth.

For narrowband systems, the impact on baseband complexity is relatively low. However, the complexity increases significantly when ‘wideband’ systems are considered, i.e., when considering operation in frequency-selective channels. However, it might well be that the potential capacity/spectrum utilisation gains would outweigh the cost of increasing the processing power of the radio.

Significant capacity gains can be achieved in a MIMO-enabled system by dynamically adjusting to the current radio environment. A non-adaptive system needs to be designed for the worst-case channel. But for an adaptive system, adaptive modulation and coding can be used to maximise the capacity of each signal path in near real-time. SDR is well suited to the implementation of adaptive MIMO because an SDR has the potential to offer the flexibility and reconfigurability required in an adaptive MIMO system.

Waveforms

Waveforms have been considered in this chapter, and it is broadly a tutorial and a literature review. Here, the term ‘waveform’ refers to the characteristics of the physical layer signal, e.g., bandwidth, symbol rate, modulation scheme, etc. Because of the different requirements of the many wireless standards in use today, numerous waveforms exist. Considering the choice of modulation scheme, a modulation method suitable for interference-tolerant, low data rate communications might be completely unsuitable for providing spectrally efficient, high-data rate communications, for which tolerance to interference might be less of an issue (e.g., due to reduced operating ranges, lower mobility or through relaxed latency requirements).

Since numerous waveforms exist, the motivation to develop SDRs exhibiting interoperability, i.e., integrating multiple waveforms, is strong. Waveform parameters that affect the ability to integrate waveforms in an SDR have been identified and the current major wireless standards have been reviewed from a waveform’s viewpoint. In addition, a literature review of integration efforts with respect to waveforms has been presented.

Integration at the RF front-end (including frequency translation) should be distinguished from integration of the baseband processing. Implementation of the RF front-end is governed mostly by the waveform’s bandwidth and band position. Its sensitivity and blocker characteristics are also important. Implementation of the baseband processing is heavily dependent on the modulation method adopted and the waveform’s complexity.

UNCLASSIFIED

Many multi-band and multi-mode solutions have already been applied to combine waveforms for mobile, wireless local and personal area networks. In some cases it has been possible to exploit similar characteristics between different waveforms to enable lower-complexity standards to be absorbed into those with higher computational requirements. However, systems attempting to integrate higher frequency bands, such as the 5 GHz HIPERLAN/2 and IEEE 802.11a, have yet to be realised using SDR platforms; such systems are either implemented using application-specific integrated circuits (ASICs), or implement only a subset of the standards being integrated. Thus, insufficient baseband processing power is currently a limiting factor. However, it is reasonable to assume that a rise in baseband processing power ultimately might stimulate the full integration of mobile and WLAN standards.

It is likely that the capabilities of the antenna system and the RF front-end will be limited at design time. This being the case, the creation of 'future-proof' SDR devices will probably be restricted to accommodating limited changes to the baseband processing. The antenna and RF front-end will limit operation to predefined frequency bands, whilst history has shown that each generation of wireless standard generally requires a significant increase in baseband processing power, i.e., simply because a platform can be reprogrammed does not necessarily mean that it may be upgraded to support all future wireless standards.

Software Aspects

Various software aspects pertinent to the implementation of SDR software have been considered, in particular, issues relating to the development of platform-independent software and the operation of SDRs in cognitive networks.

A review of existing approaches to the development of platform-independent models reveals that useful levels of abstraction that might be applied to the development of SDR software are possible. Additional aspects that might be important to the development of SDR software are provided by the detailed models defined as part of the web services architecture, developed by the World Wide Web Consortium. In particular, the message-orientated model, the service-orientated model, the resource-oriented model and the policy model are all relevant.

Whilst these models might be used to develop platforms supporting runtime flexibility, for example, it is important that such flexibility is not exploited in an uncontrolled manner. Policies are proposed as a possible means of controlling adverse behaviour emerging in SDR networks.

It is likely that some of the principles of grid computing (i.e., the controlled sharing of resources in conjunction with parallel and distributed computing) might be applied to SDRs to enable networks of SDR nodes to share and better manage their capabilities. Techniques such as semantically rich peer-to-peer resource trading and policies might be used to tackle issues such as self-interest and mixed trust. These are all current areas of research. It has been concluded that commercial and legislative issues, rather than technical issues, are likely to pose the greater obstacles to the realisation of cognitive SDR grids.

Security

Security is a key issue in any wireless communications system. However, SDR introduces some unique security challenges, especially when considering SDRs that can be reconfigured across the air interface. This part of the study focuses on the security issues that are unique to SDR.

In accordance with the long-term vision of SDR, SDR devices might be readily upgraded to operate with different waveforms. This has a number of potential benefits. For example, it could help improve information security as it gives the radio the ability to switch to more robust waveforms as and when required. Furthermore, due to the ease of upgrading units, it would be relatively straightforward to ensure that the devices were kept up to date with the latest security developments. An ability to download new software will also enable the latest standards to be adopted to help optimise spectrum utilisation.

However, if SDR devices can be modified remotely, concerns are raised about securing the exchange of operational information. If a software download were intercepted, an intruder might configure his/her terminal with the intercepted software and masquerade as the intended user. Thus, the intruder might be able to listen in on transmissions, download software onto unauthorised hardware platforms or make use of services for which the intended user was being charged.

Moreover, if an intruder were able to intercept software downloads, he/she might be able to modify them to change the nominal operating parameters, e.g., to ‘snoop’ on other traffic, circumnavigate billing mechanisms or deny service to others. To prevent a situation like that, user authentication is a key issue in SDR security. An associated issue is the authorisation of hardware/software pairs to ensure that unauthorised hardware/software pairings are prevented from being used. Thus, a certification role for the regulatory bodies needs to be considered.

It has been concluded that, whilst an ability to download new software might be used to help better protect a user’s data or help optimise spectrum usage, it will be essential that every effort is made to protect and authenticate the download process itself. Similarly, it will be crucial that any ‘control’ data are protected. SDR and CR research raises the idea of using spectrum more efficiently through a dynamic spectrum market. This raises security concerns, as dynamic spectrum trading infers that there will be billing and payment information in transit. If these data are intercepted, an intruder might make use of spectrum allocated to, and paid for by another user.

Technologies already exist to enable the secure implementation of SDRs. However, standardisation will be a significant obstacle; it is essential that procedures are standardised to ensure compatibility between equipment and to help simplify the potentially daunting task of managing SDR software.

Finally, the security implications of using ‘traditional’, general purpose processors in the implementation of SDRs have been considered. In particular, there are concerns regarding the use of ‘shared’ memory in such processor architectures. Without proper safeguards, such architectures are vulnerable to malicious attack and accidental malfunction resulting from programs being able to access memory allocated to other processes. Several technologies exist that might enhance the security architecture of an SDR. These include the use of ‘virtual machines’, resorting to physically separate hardware or the use of a memory control module to manage memory access.

Cognitive Radio and Radio Management

The aim of radio management is to ensure that the scarce resource of RF spectrum is allocated amongst its multiple and diverse range of users in a manner as efficient as possible. Under the current centralised and static ‘command and control’ method of spectrum allocation, there are large amounts of spectrum that are ‘wasted’. This is because there are bands of allocated spectrum that are not used continuously, or are used only in specific geographic regions.

The flexibility of SDR might be used to promote more efficient spectrum usage. In particular, SDR is likely to be an enabling technology for cognitive radio (CR). CRs may have an awareness of the propagation environment in which they are operating, know or be able to anticipate the level of service the user requires, or know the geographic location of the device. This knowledge would enable a reasoned decision to be reached about the best and most efficient waveform to be used at any time. As with current devices, the spectrum requirement of CR is not constant, but the key difference is that a CR has an awareness of its spectrum requirement at any time and could report this to other users. In this way CR could provide the opportunity for dynamic spectrum allocation (DSA) and trading, in which all radio devices know their own spectrum requirement and negotiate to fit in with the spectrum currently available.

We have concluded that, there are a number of generic issues that will need to be addressed regarding DSA:

- Radio etiquette: Formalising the rules of spectrum trading, determining who will have priority over a particular portion of the spectrum and defining what constitutes unauthorised use of the spectrum.
- Billing: Determining how to reliably bill users of spectrum resources in an open and dynamic market.
- Policing: There will be a need to monitor the spectrum for illegal usage and to define procedures to tackle persistent offenders.
- Quality of service: If a device must compete for spectrum, then there may be occasions when the user will not obtain the quality of service he/she expects because available bandwidth cannot necessarily be guaranteed.
- Increased overhead: In a dynamic spectrum market there will be an additional overhead in network traffic due to negotiations taking place to secure spectrum rights. The effect of this additional overhead will need to be investigated.

CR has many potential benefits. For example, producing a radio that can participate in dynamic spectrum trading is probably a long way off. However, a radio that monitors the received signal-to-noise ratio and decides whether or not the overhead of forward error correction (FEC), for example, is necessary might be much more feasible. Similarly, there might be gains to be made by using CRs that dynamically switch modulation scheme to reach an optimal compromise between signal bandwidth, waveform resilience and data throughput. Finally, we have noted that, under the current regulatory environment, the use of CR and spectrum trading to improve the efficiency with which the available radio spectrum is utilised is prohibited due to the existence of internationally recognised, harmonised frequency bands.

UNCLASSIFIED

Regulatory Issues

SDR promises to revolutionise the design and operation of radio equipment and its potential advantages are many. However, SDR might also bring with it a number of regulatory difficulties, especially when highly flexible, generic radio platforms supporting reprogramming via field updates are considered.

For the purposes of this study SDR is considered to be applicable to the bottom two layers of the seven-layer open standards interface (OSI) reference model, the physical and data link layers, together with some of the functionality of layer three, the network layer. Thus, in an SDR, the characteristics of the transmitted waveform are defined in software, as are the structure and formatting of the physical and logical channels. Changes to the software have the potential to modify the spectrum of the transmitted signal and the interoperability with other users.

While it is reasonable to assume that SDRs are reprogrammable, in the context of regulatory issues, two classes of SDR have specific relevance. These are SDRs which are programmed at the time of manufacture and cannot be reprogrammed once released to the market, and SDRs that can be reprogrammed in the field, either via an over-the-air update or by some other means. Considering non-reprogrammable radios, SDR simply represents an alternative method of implementation over that for more traditional radio architectures. As such, there are no regulatory issues specific to SDR. Regulatory issues posed by SDR are specific to radios that can be reprogrammed after manufacture. One solution would be to refuse approval of radio equipment that is reprogrammable. However, such a response would prevent the full potential of SDR from being exploited and inhibit the development of innovative radio systems that might facilitate the efficient use of the available RF spectrum.

The main issue regarding the regulation of field reprogrammable SDRs is that of type approval. In the United States, the Federal Communications Commission (FCC) has already undertaken various studies into SDR and has reviewed its regulatory processes as a result. In particular, the FCC has defined a new class of permissive change that allows the fast-tracking of software updates. Thus, software updates can be released without having to undergo again the entire type approval process. Furthermore, the FCC is to allow the electronic display of relevant FCC authorisation numbers in place of the physical equipment marking. This relaxation theoretically makes it possible for SDRs to be reprogrammed in the field with completely new configuration data without requiring the return of the hardware to authorised service centres.

Despite these concessions, the FCC type approval process is still quite restrictive. Software updates are only to be permitted following the testing of all software/hardware combinations. Furthermore, streamlined software updates may only be issued by the original authorisation applicant. Therefore, if a third party wishes to develop software to operate on an original equipment manufacturer (OEM) radio platform, they must apply for full type approval of the software/hardware combination, that is, they must take responsibility for the conformance of the hardware as well as their software.

UNCLASSIFIED

In the UK the type approval of radio equipment is controlled by the R&TTE Directive 1999/5/EC. Under this directive the manufacturer is responsible for the conformance of its products. Furthermore, the manufacturer must give 28 days notice before releasing products that operate non-harmonised radio technologies and/or in non-harmonised frequency bands. The Directive does not appear to be particularly compatible with the concept of field-programmable SDR equipment, especially when considering the scenario in which one party manufactures the hardware and others develop the software. It is not clear who must take responsibility for the conformance of the hardware; without the software, the hardware need not be specific to any particular radio technology. However, software houses are unlikely to want to take responsibility for the hardware once their software is deployed.

SDR is likely to become an enabling technology in the adoption and acceptance of the concepts of spectrum trading and spectrum liberalisation. Ofcom is keen to promote these concepts with a view to improving spectrum management and encouraging users to make better use of the spectrum available to them. Once spectrum trading and liberalisation have become the accepted norm, advanced CR systems are likely to emerge, which will further exploit the flexibility and reconfigurability of SDR. However, the existence of harmonised frequency bands might be an obstacle to realising the full potential of spectrum liberalisation and CRs.

Commercial Drivers

The commercial drivers for the manufacturer, the network operator and the user for the implementation of SDR have been discussed – both the potential advantages and disadvantages.

The strongest drive towards SDR would appear to be by the network operators. The observed advantages include the prospect of more ‘future-proof’ equipment and the potential ability to rapidly introduce new technologies and the latest services.

Manufacturers of network infrastructure equipment also seem to be keen on adopting SDR technology. Adopting SDR technology means that potentially they can get products to market faster, and continue development, fixing bugs and adding new features, post-manufacture. This is especially true in the case of emerging wireless technologies. Here, product development can be started earlier and continued in parallel with the finalisation of the wireless specifications. Thus, not only is risk reduced, but also there is a potential for reduced-functionality hardware to be made available earlier for the purpose of test-bed trials. Test equipment manufacturers have been exploiting SDR techniques for many years. Rather than developing dedicated hardware for each and every application, a streamlined range of generic test equipment is developed and then targeted at the appropriate applications through the development of customised software. Not only does this allow hardware development costs to be spread, it also helps to improve customer ‘loyalty’.

The manufacturing group with perhaps the weakest drive towards SDR is that of the handset manufacturers. The factors that tend to drive the design of the latest handsets (e.g., size, weight, power consumption) generally conflict with the characteristics of current SDR technologies. Power consumption, in particular, is likely to be a significant hurdle to the widespread adoption of SDR technology in handset design. The implication is that, if SDR is to become an enabling technology for the concepts of CR, spectrum trading and spectrum liberalisation, then the factors preventing the development of SDR handsets will need to be addressed.

UNCLASSIFIED

Finally, considering SDR from the perspective of the user, the factors that tend to influence handset choice are often restricted to cosmetic appearance, size, weight, battery life and feature set. Ultimately, SDR technology might enable multiple wireless standards to be integrated into the handset. Beyond this, however, the user is unlikely to choose one handset over another simply because it incorporates SDR technology. On the contrary, the likely side effect of reduced battery life and increased size and weight (i.e., resulting from the need for a larger battery) might even be a deterrent. Therefore, until wireless systems are pushed to the point where the flexibility offered by SDR becomes a requirement, there is unlikely to be much demand by the user to see the emergence of SDR handsets.

Assessment of SDR's First Applications and Areas of Deployment

The commercial applications of SDR have been considered in this chapter; the potential deployment of SDR in cellular networks, emerging wireless networks e.g. WiFi and WiMAX, transportation and public safety communications. An examination of the cellular network industry reveals how SDR has already been making its way into this sector. This has been via branded SDR products from companies like Vanu and AirNet or more discretely via the adoption of reconfigurable devices by many manufacturers into 3G base stations. The deployment of SDR base stations will be most relevant to the US due to their adoption of multiple cellular standards.

A key commercial barrier to widespread SDR deployment in this market is the lack of interoperability with existing network infrastructures, arising from the initial uncoordinated efforts by various organisations. Standardisation initiatives like the Open Base Station Architecture Initiative (OBSAI) could help to tackle the interoperability issue.

Emerging wireless network services like WiFi and WiMAX are being seen as additions to existing cellular service portfolios. This could generate the requirement for a mobile device to be able to roam between services, which could in turn boost SDR. Power constraints mean that it will be some years before SDR is implemented in handsets. However, when this does become feasible it could trigger the widespread adoption of SDR because it would give the user multiple services in a single device.

Outside the cellular industry, public safety and transportation are two areas that should not be neglected by SDR developers. The public safety sector, particularly in the US, is facing an issue with a lack of interoperability across services. This is similar to the situation faced in the defence sector which led to the Joint Tactical Radio System (JTRS) programme. SDR could offer an attractive solution to this problem. Long procurement cycles mean that the adoption of SDR into this market will be slow. However, investment in a public safety version of the JTRS could give real incentives for SDR developments in this area.

Finally, interest in Intelligent Transport Systems (ITS) in the transportation industry could provide a good opportunity for SDR to dominate an emerging market from the outset. SDR could provide the solution for deploying ITS that are interoperable across national and international borders. This is an emerging market that will take time to define itself but also one that SDR developers should ensure they are actively involved in from the outset.

UNCLASSIFIED

In the long-term, SDR fits with the future direction of communications as the industry moves towards more complex and flexible systems incorporating more services. SDR will also be key in supporting movement towards more spectrally efficient schemes and eventually dynamic spectrum trading. It is a key enabling-technology for Cognitive Radio, which in turn supports visions of flexible spectrally aware devices.

A forecast of SDR roll-out has been provided for various commercial wireless technologies, and it has been noted that the accuracy of this forecast is dependent upon a range of parameters including the success of initiatives like OBSAI, partnerships between SDR start-ups and main manufacturers and future regulation of SDR devices.

Spectrum Efficiency Gains of SDR and CR

Spectrum efficiency gains of SDR and CR have been considered in this chapter. The move to increase the spectrum available for unlicensed operations (to enable greater scope within the frequency spectrum for CR) will permit innovative, short-range radio technologies to be developed. Such technologies might be used at home and in the workplace to provide low-cost wireless connectivity and relieve the load on longer-range cellular systems.

An application of CR that has received significant interest is that of dynamic spectrum allocation (DSA). We can use analysis of the Erlang B formula to investigate how DSA might improve spectral efficiency. Consider the mean channel utilisation as a function of the number of physical channels, for a blocking probability of 2%. Let us assume that we have 100 channels to be distributed between five operators. If we allocate the available channels equally, each operator has 20 channels. We find that mean channel utilisation is approximately 65%. With DSA, all five operators have access to all 100 channels. Here we find that mean channel utilisation is increased to approximately 86%. In other words, using DSA in our example instead of fixed spectrum allocation theoretically increases overall capacity by approximately 33%. Therefore, the channel utilisation benefits available from DSA provide further motivation for the development and application of SDR.

Two approaches to DSA have been considered. In the first approach, a ‘broker’ manages a central database to control access to licensed spectrum. This approach is best suited to client/server, i.e., cellular systems, with the fixed infrastructure handling the acquisition of temporary spectrum access rights. Another form of DSA is that of opportunistic DSA, nominally in unlicensed spectrum. Here, devices use detection methods to search for and use ‘unused’ spectrum. There are many technical challenges associated with opportunistic DSA, not least how to ensure, with a high probability of success, that interference is not inadvertently caused to nearby legacy spectrum users that remain undetected by the CR.

Finally, the use of dynamic waveform configuration to maximise system capacity has been considered. Whereas DSA aims to ‘pack’ users into unused spectrum, dynamic waveform configuration actually attempts to improve the spectral efficiency of a given system directly. Two forms of dynamic waveform configuration have been considered, namely, closed-loop power control and dynamic modulation and/or channel coding. Both methods use the concept of a target signal-to-noise ratio (SNR).

UNCLASSIFIED

With dynamic power control, excess SNR allows the transmit power to be reduced. Thus, interference caused to other co-channel users can be reduced. The net result of this is that the separation between different systems sharing common spectrum may be reduced (yielding improved spectrum utilisation); this technique is already implemented in current cellular systems.

With dynamic modulation and/or channel coding, excess SNR is used to support higher-order modulation methods and/or a decreased forward error correction (FEC) coding rate. By increasing the channel throughput for users operating close to the transmitter, additional users may be supported in the same spectrum, thereby increasing system capacity, again current cellular systems employ variants of this technique.

In conclusion, it has been found that techniques such as DSA and dynamic waveform configuration have significant potential to improve the spectral efficiency of future radio systems. Multi-mode terminals that take advantage of local-area access points also have the potential to improve overall spectral efficiency by relieving the load on congested cellular networks in traffic hotspots. The inherent flexibility of SDR is well suited to the implementation of these approaches to improved spectral efficiency and represents an ideal platform for the realisation of high-performance CRs. Therefore, SDR in conjunction with CR will provide the optimal utilisation of the radio spectrum

Summary

Many topics associated with SDR have been considered within this study. The result is an informative reference document providing a broad background on some of the key issues associated with SDR and CR. An assessment of SDR including potential obstacles, enablers and the spectrum efficiency gains of SDR has been presented in this document.

It is concluded that the emergence of SDR as a practical, realisable technology will mark a significant milestone in the evolution of radio. Throughout the document, examples have been provided of the current development of SDR, in connection with both specific components within an SDR, as well as, in connection with current wireless communication standards. It is important to recognise that SDR, as a technology, is still in its infancy. Ultimately, however, SDR promises to open the door to concepts such as CR and the development of novel, spectrally efficient wireless communication systems.

As SDR technology is maturing over the coming years, practical reality will shake off some of the more blue-sky, long-term, idealised visions of SDR. From a regulatory standpoint, it is essential that the development of SDR is reviewed regularly, for the foreseeable future at least, because overly restrictive regulation is likely to prohibit innovation and thus be counterproductive. By the same token it might be prudent initially to be cautious, and use the knowledge and experience of practical systems to shape the regulatory approach long-term.

Epoch estimations related to SDR have been provided in the relevant sections of this document for the course of the next ten years, or so. However, the forecasts presented in this study may contain some deviation as these forecasts are strongly influenced by many factors, ranging from changes in market demands to the discovery of new technologies, and such developments may therefore severely distort these predictions.

List of contents

1	Introduction	25
1.1	Relation between the Main Document and the Overview Document	25
1.2	Contractual Information	25
1.3	Aim	25
1.4	The SDR Concept	26
1.4.1	Open System Interconnection (OSI) reference model and SDR	26
1.4.2	SDR Definition from The SDR Forum	27
1.4.3	Additional views on the definition of SDR	28
1.5	Cognitive Radio	30
1.6	Background	31
1.7	Consortium Members	32
1.8	Report Structure	34
1.9	Reference	36
1.10	Acknowledgements	36
2	Antennas	37
2.1	Introduction	37
2.2	Software Defined Radio	37
2.2.1	Key Components	38
2.2.2	SDR Antenna Requirements	40
2.3	Antenna Fundamentals	40
2.4	Antenna Options for SDR	43
2.4.1	Phased Array Antennas	44
2.4.2	Reconfigurable Antennas	47
2.4.3	Antennas for Polarisation Diversity	50
2.4.4	Multi-Band and Ultra-Wideband Antennas	54
2.4.5	Plasma Antennas	58
2.4.6	Antenna Modelling	59
2.5	Epoch for the Development of Antenna Technology	60
2.6	Conclusions	61
3	Radio Frequency (RF) Linearisation	62
3.1	Introduction	62
3.1.1	Software Defined Radio	62
3.2	RF Linearisation and SDR	63
3.3	RF Linearisation and Mobile Communication Standards	64
3.3.1	Power Amplifier Linearisation in 2G Systems	64
3.3.2	Linearisation in 3G and Beyond	64
3.4	Power Amplifier Linearisation Methods	65
3.4.1	Analogue Methods	65
3.4.2	Digital Methods	67

UNCLASSIFIED

3.5	Technology	75
3.5.1	ADC/DAC Performance Overview	75
3.5.2	Digital Signal Processing Hardware in SDRs	76
3.6	Epochs for the Developments of RF Linearisation Technology	81
3.7	Conclusions	82
4	Antenna Processing	83
4.1	Introduction	83
4.2	Receiver and Transmitter Architectures	83
4.2.1	Receiver	84
4.2.2	Transmitter	86
4.3	Analogue-to-Digital and Digital-to-Analogue Conversion	87
4.3.1	Sampling	87
4.3.2	Quantisation	91
4.3.3	Quality Measures	93
4.3.4	Requirements	93
4.4	Commercially Applied Technology and Available Devices	94
4.4.1	ADC and DAC Conversion Techniques	95
4.4.2	Commercial ADC and DAC Devices	98
4.5	Research Development and Future Devices	100
4.5.1	Time-Interleaved ADCs	101
4.5.2	Optical Sampling	101
4.5.3	Superconducting Devices	102
4.6	Epochs for the Developments of RF Antenna Processing Technology	102
4.7	Conclusions	103
5	MIMO Technology and SDR	104
5.1	Introduction	104
5.2	Introduction to MIMO Techniques	104
5.2.1	MIMO Channel Models	108
5.3	Literature Review	110
5.3.1	SDR Implementations of MIMO Systems	110
5.3.2	MIMO in Existing and Proposed Standards	111
5.4	RF Architectures for MIMO Implementation in SDR	118
5.4.1	Antennas	118
5.4.2	RF Hardware: Separate Implementation	124
5.4.3	Analogue RF Implementation	124
5.4.4	Multiplexed RF Chains	125
5.5	Outline of MIMO SDR Implementation	126
5.5.1	STBC	126
5.5.2	Spatial Multiplexing	127
5.5.3	Wideband Systems	128
5.6	Adaptive MIMO and SDR	136
5.6.1	Adaptive Modulation and Coding	136

UNCLASSIFIED

5.6.2	Receiver Beamforming	138
5.6.3	MIMO Exploiting Channel Knowledge at the Transmitter	140
5.6.4	On the Issue of MIMO Implementation for SDR	148
5.7	Conclusions	149
6	Waveforms	151
6.1	Introduction	151
6.2	Factors for the Integration of Waveforms	151
6.2.1	RF Front-End and Up- and Down- Conversion	152
6.2.2	Baseband Processing	153
6.3	Wireless Standards	154
6.3.1	Mobile Standards	154
6.3.2	Wireless Local Area Networks	156
6.3.3	Wireless Personal Area Networks	158
6.3.4	Wireless Metropolitan Area Networks	159
6.4	Wireless Communications	161
6.4.1	Amplitude Shift Keying	161
6.4.2	Phase Shift Keying	161
6.4.3	Quadrature Amplitude Modulation	162
6.4.4	Orthogonal Frequency Division Multiplexing	163
6.4.5	Direct Sequence Spread Spectrum	163
6.4.6	Frequency Hopping Spread Spectrum	164
6.4.7	Gaussian Frequency Shift Keying	164
6.5	Efforts at Integrating some Wireless Standards	165
6.5.1	Single Band Systems for Mobile Standards	165
6.5.2	Systems Integrating 2G and 3G Wireless Mobile Standards	166
6.5.3	Systems Integrating 3G and WLAN	169
6.5.4	WLAN and WPAN Integration	170
6.5.5	Further Integration Issues	170
6.6	Conclusions	171
7	Software Aspects	173
7.1	Introduction	173
7.2	Open Software Architectures	174
7.2.1	Motivation for the Software Communications Architecture	174
7.2.2	Development of the Platform-Independent Model for SCA	176
7.2.3	Reference Implementation of SCA	179
7.3	Technologies towards Cognitive Networks	180
7.3.1	Managing Faults in SDR Networks	180
7.3.2	Controlling SDR Interactions through Policy Based Control	183
7.3.3	SDR Applicable Techniques from the World Wide Web Consortium (W3C)	185
7.3.4	SDR and Grid Computing	189
7.4	SDR and Human Interface	192

UNCLASSIFIED

7.4.1	Introduction	192
7.4.2	Identified Users	192
7.4.3	HMI for SDR Equipment	192
7.4.4	HMI for SDR Networks	193
7.4.5	HMI for SDR Development	193
7.5	Conclusions	194
8	Security	196
8.1	Introduction	196
8.2	What is Security?	196
8.2.1	Authentication	196
8.2.2	Integrity	197
8.2.3	Confidentiality	198
8.2.4	Availability	198
8.2.5	Summary	199
8.3	Security Analysis of an SDR Environment	200
8.3.1	The Threat Environment	201
8.3.2	Types of Traffic in an SDR Environment	202
8.3.3	The Requirements of a Secure SDR System	203
8.3.4	Summary	208
8.4	Hardware Security	209
8.4.1	The Insecurity of a General Purpose Processor	209
8.4.2	Technologies for Providing Memory Security	210
8.5	Conclusions	212
9	Radio Management and CR	213
9.1	Introduction	213
9.2	The Current Radio Management Scheme	214
9.3	Spectrum Trading	216
9.3.1	Permanent to Semi-Permanent Transfer of Spectrum Allocation	216
9.3.2	Long-Term Dynamic Spectrum Allocation	216
9.3.3	Short-Term Dynamic Spectrum Allocation	217
9.4	Cognitive Radio	222
9.4.1	Definitions of Cognitive Radio	222
9.4.2	Possibilities for Cognitive Radio	223
9.4.3	Examples of Primitive Cognitive Life Forms	225
9.5	Technical Barrier and Enabling Technologies	226
9.5.1	Radio Etiquette	227
9.5.2	Legacy User Protection and Coexistence	227
9.5.3	Spectrum Trading	228
9.5.4	Regulatory Issues	228
9.6	Conclusions	229
10	Regulatory Issues	230
10.1	Introduction	230

UNCLASSIFIED

10.2	What is SDR and at What Point Does the SDR Part Stop?	230
10.3	What are the Benefits of SDR?	233
10.3.1	The Manufacturer	233
10.3.2	The Regulator	234
10.3.3	The Network Operator	234
10.3.4	The User	235
10.4	Why is SDR Such a Potential Regulatory Minefield?	235
10.5	Regulatory Aspects for Different SDR Applications	236
10.5.1	Base Station SDRs	237
10.5.2	User Equipment SDRs with Preloaded Configuration Data	238
10.5.3	Reconfigurable User Equipment SDRs for Use in Client/Server Networks	238
10.5.4	Reconfigurable User Equipment SDRs for Use in Ad-Hoc, Peer-to-Peer Networks	240
10.5.5	Reconfigurable User Equipment SDRs for Use in Cognitive Radio Networks	240
10.6	Type Approving SDR Applications	241
10.7	Equipment Type Approval in the United States by the FCC	243
10.8	Equipment Type Approval within the EU and the R&TTE Directive	245
10.9	SDR and Spectrum Trading and Liberalisation	246
10.10	Conclusions	249
11	Commercial Drivers	251
11.1	Introduction	251
11.2	A Manufacturer's Perspective	251
11.2.1	Infrastructure Vendors	251
11.2.2	Terminal Manufacturers	255
11.2.3	Semiconductor Houses	256
11.2.4	Test Equipment Manufacturers	258
11.2.5	Software Houses	258
11.3	A Network Operator's Perspective	259
11.3.1	Commercial Drivers for the use of Programmable Devices in Equipment	259
11.3.2	Commercial Drivers for Post-Manufacture Reconfigurability	260
11.3.3	Commercial Drivers for Reconfigurable Terminals	261
11.4	A Consumer's Perspective	262
11.5	Conclusions	262
12	Assessment of SDR's First Applications and Areas of Deployment	263
12.1	Introduction	263
12.2	SDR in Cellular Networks	263
12.2.1	Standards Evolution in Cellular Networks	264
12.2.2	Current SDR Cellular Network Products	266
12.2.3	3G Roll-out and SDR	271
12.2.4	Mobile Phones	275
12.2.5	Summary	276

UNCLASSIFIED

12.3	SDR in other Commercial Applications	277
12.3.1	WiFi, Bluetooth and WiMax	277
12.3.2	GPS	279
12.3.3	PDAs	280
12.3.4	Transportation: Automotive and Commercial	281
12.3.5	Summary	282
12.4	SDR in Public Safety Mobile Radios	283
12.4.1	The Need for Harmonisation across the Emergency Services	283
12.4.2	Requirements of Public Safety Mobile Networks in Contrast to Commercial Cellular Networks	284
12.4.3	Public Safety Market within the UK	285
12.4.4	Emerging SDR Products within Public Safety	285
12.4.5	Summary	286
12.5	Forecast of SDR Roll Out	287
12.5.1	Future Benefits of SDR	287
12.5.2	Comparison of SDR Popularity across Market Sectors	289
12.5.3	Predicted Trends for SDR Deployment	293
12.6	Conclusions	294
13	Spectrum Efficiency Gains of Software Defined Radio and Cognitive Radio	296
13.1	Introduction	296
13.2	Existing Systems Exhibiting Limited Cognitive Capabilities	296
13.2.1	Digital Enhanced Cordless Telecommunications	296
13.2.2	Universal Mobile Telecommunications System (UMTS)	297
13.3	Multi-Mode Terminals	299
13.4	Dynamic Spectrum Allocation	302
13.4.1	How Can DSA Help Improve Spectrum Utilisation?	302
13.4.2	Brokered DSA in Licensed Spectrum	307
13.4.3	Opportunistic DSA in Unlicensed Spectrum	310
13.4.4	Opportunistic DSA in Licensed Spectrum	311
13.5	Dynamic Power Control, Modulation and Channel Coding	312
13.5.1	Dynamic Power Control	312
13.5.2	Dynamic Modulation and/or Channel Coding	315
13.6	Conclusions	319
14	Conclusions	321
14.1	Epoch Estimation for Technology Development	322
14.2	General Conclusions	323
14.3	Summary	324

UNCLASSIFIED

References / Bibliography	325
Abbreviations	348
A Abstracts from SDR and MIMO Search	356
B Selected Google Search Results for SDR and MIMO	360
C Web Services Security Specifications	364

1 Introduction

1.1 Relation between the Main Document and the Overview Document

This Main Document is a detailed companion to the Overview Document [1], which provides a concise account of an evaluation of Software Defined Radio (SDR). It is the intention of this document to provide a detailed account on the various subjects contained with an evaluation of SDR and to allow numerous references to be readily identified.

1.2 Contractual Information

This document has been produced by QinetiQ, Defence and Technology Systems for Ofcom under contract number 410000262 and provides an evaluation of software defined radio.

1.3 Aim

The aim of the work reported in this document is to provide an evaluation of software defined radio (SDR). In broad terms, the following three key areas are addressed:

- Technology Requirement Enablers
- Assessment of SDR
- Spectrum efficiency gains of SDR and cognitive radio (CR).

In the context of subject categories, the following key areas are outlined within this study:

- Antennas
- RF (Radio Frequency) Linearisation
- Antenna Processing
- MIMO (Multiple-Input, Multiple-Output) Technology and SDR
- Waveforms
- Software Aspects
- Security
- Radio Management and CR
- Regulatory Issues
- Commercial Drivers
- An Assessment of SDR's First Applications and Areas of Deployment
- Spectrum Efficiency Gains of SDR and CR.

1.4

The SDR Concept

The concept of an SDR has evolved over the past three or four decades, with applicability to various parts of a radio. Over this period, a wide range of sources have put forward their definition of SDR. This gives rise to various interpretations of what SDR actually is, and what SDR is not. A key reason for the degree of variability in the perception of SDR is that SDR is an all-embracing term that may, quite rightly, be applied to a wide range of radio platforms. The differing visions arise from their authors focusing on particular attributes of the ‘ultimate’ SDR. We start by briefly discussing what we mean by SDR.

Perhaps the single most important attribute of any SDR is that its behaviour at the physical layer (which is considered in more detail in section 1.4.1), e.g., transmit power, frequency, modulation scheme, etc., may be changed in software without requiring any additional hardware modifications.

1.4.1

Open System Interconnection (OSI) reference model and SDR

It is important that SDR is not confused with application software and other software not associated with the radio. As stated previously, SDR describes the software emulating part, or all, of the signal path. Thus, considering the OSI reference model, shown in Figure 1-1, SDR refers in general to functionality within the physical and data link layers and perhaps parts of the network layer. Functionality in the higher layers is not specific to SDR and should not therefore be classed as such.

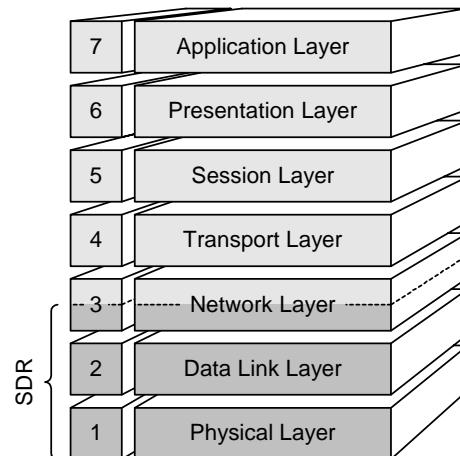


Figure 1-1: The seven-layer OSI reference model

For the purposes of this section, we will assume the following definitions for the terms ‘software defined radio’ and ‘radio’.

It is reasonable to assume that SDRs are “radios that provide software control of a variety of modulation techniques, wideband or narrowband operation, communications security functions (such as frequency hopping), and waveform requirements of current and evolving standards over a broad frequency range”.

In addition, the term ‘radio’ is taken to mean “any system that seeks to communicate with another system by means of a modulated signal within the RF spectrum”. Typical examples might be a mobile phone, a personal digital assistant (PDA), a mobile radio or a wireless-enabled laptop computer.

UNCLASSIFIED

1.4.2 SDR Definition from The SDR Forum

The SDR Forum, a non-profit corporation set up to support the development, deployment and use of open architectures for advanced wireless systems, have developed a multi-tiered definition of SDR. The five-tier concept is summarised in Table 1-1. Tier Zero represents ‘traditional’ radio hardware and forms a baseline reference. The uppermost tier, Tier Four, represents the ‘ultimate’ vision of SDR. Reality falls somewhere in the middle. For most applications, state-of-the-art SDR currently aligns with the Tier Two definition. Note that, as mentioned above, it may be argued that virtually all modern wireless communications equipment may be classified as being software-controlled radios (i.e., Tier One).

Tier	Name	Description
0	Hardware radio (HR)	Baseline radio with fixed functionality.
1	Software-controlled radio (SCR)	The radio’s signal path is implemented using application-specific hardware, i.e., the signal path is essentially fixed. A software interface may allow certain parameters, e.g., transmit power, frequency, etc., to be changed in software.
2	Software defined radio (SDR)	Much of the waveform, e.g., frequency, modulation/demodulation, security, etc., is performed in software. Thus, the signal path can, with reason, be reconfigured in software without requiring any hardware modifications. For the foreseeable future, the frequency bands supported may be constrained by the RF front-end.
3	Ideal software radio (ISR)	Compared to a ‘standard’ SDR, an ISR implements much more of the signal path in the digital domain. Ultimately, programmability extends to the entire system with analogue/digital conversion only at the antenna, speaker and microphones.
4	Ultimate software radio (USR)	The USR represents the ‘blue-sky’ vision of SDR. It accepts fully programmable traffic and control information, supports operation over a broad range of frequencies and can switch from one air-interface/application to another in milliseconds.

Table 1-1: SDR Forum’s tier definitions [2]

UNCLASSIFIED

In order to aid the reader in visualising the concepts identified in Table 1-1, it is possible to consider the functionality contained within a radio and to identify, in broad terms, how the concepts presented by the SDR Forum's five-tier definition fit within a radio. Figure 1-2 illustrates an abstraction of the five-tier definition, where the length of the arrow indicates the proportion of the software content within the radio. For example, it can be seen that, at tier 0 we have very little software element by virtue of the length of the arrow being minimal. Conversely, at tier 4 we have the ultimate software radio (USR) where the entire signal at the output of the antenna has been digitised and operates within a software environment.

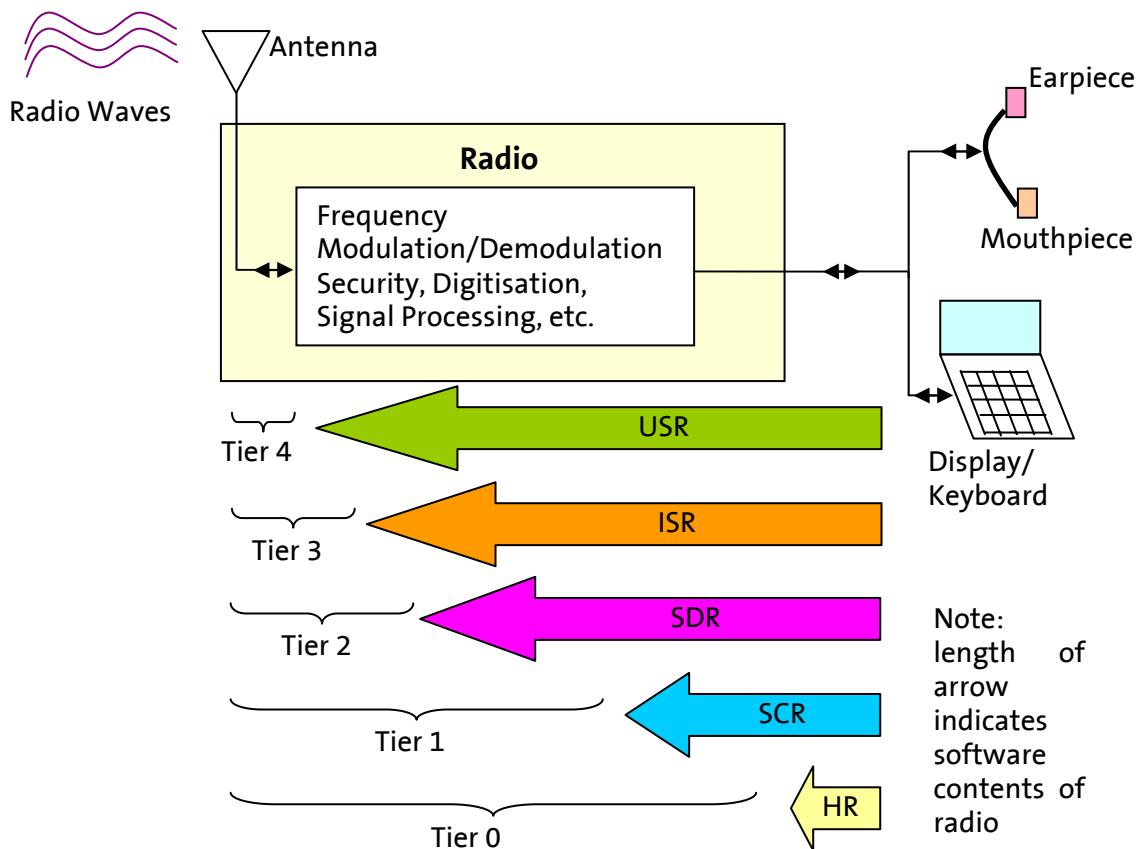


Figure 1-2: High-level Abstraction of the SDR Forum tier definition

1.4.3 Additional views on the definition of SDR

The approach adopted by the United States' Federal Communications Commission (FCC) in their definition of SDR [3] includes a requirement that the radio has an ability to operate over a very large frequency range, of e.g., 2 MHz to 2 GHz and above. This requirement, often associated with the US Joint Tactical Radio System (JTRS) programme, relates to a high-level capability.

UNCLASSIFIED

The JTRS bandwidth is likely to be too challenging for SDR, at least in the short term, and a more realistic bandwidth, over the following ten years, or so, is a frequency range of 20 MHz to 2GHz, which is typically assumed to be the appropriate bandwidth for the purposes of this study.

SDR: High-level and low-level Operation

Some definitions of SDR consider its high-level operation and simply require that certain waveform attributes can be modified under the control of software. Here, the radio may be treated as a ‘black box’ with some form of software interface to allow parameters such as frequency and output power to be controlled. Conversely, some definitions consider its operation at a much lower level and focus on how the physical layer processing is implemented, i.e., requiring that it is, in part at least, performed in software or in reprogrammable devices such as field-programmable gate arrays (FPGAs). The first definition describes software-configured radio and, while it encompasses SDR, arguably applies to almost all modern radio equipment. However, the latter definition is much more prescriptive, clearly describing what is meant by the term SDR.

SDR: Field-reprogrammability

Other definitions require the radio to be field-reprogrammable, i.e., it can be reprogrammed after initial deployment. In its simplest form, this may be achieved by downloading new software over the Internet, uploading new firmware from some form of mass-storage media, or employing field engineers to reprogram the hardware with specialist equipment. Ultimately, software updates may be performed ‘over-the-air’ and without any intervention from the user. If we follow the argument that an SDR is an SDR because of how it is implemented, field-reprogrammability is not a mandatory requirement. Certainly, by definition, the function of an SDR may be altered by changing the firmware.

However, for many potential applications, there is not necessarily any reason why a user need be given a method to update this firmware after the radio has left the production line. Radios without such a programming capability are still SDRs, albeit with nominally fixed functionality. For example, an FPGA might be configured (using a configuration programmable read-only memory (PROM)) to emulate a stereo FM radio receiver with a radio data service (RDS) demodulator using sample data acquired by sampling the RF input directly. This is certainly an example of SDR, arising by virtue of the robustness to software modification within the front-end of the radio. We note that history tells us that, even without easily accessible programming ports, consumer SDRs are likely to become tempting targets for hackers (malicious or otherwise). Take, for example, the bypassing of regional encoding on DVD players.

SDR: Dynamic Changes in Functionality

SDR is sometimes discussed in the context of a radio that can modify its operation dynamically to deliver maximum perceived performance to the user and optimal spectral efficiency. More often than not such discussions are actually about cognitive radio (CR) rather than SDR. The two terms are occasionally, and incorrectly, interchanged. It is true that SDR may play an important role in the realisation of CR. However, SDR simply represents a very flexible, generic radio platform. The intelligence required to realise a CR resides at a higher level. CR is discussed in greater detail in the following section.

Looking towards the future, SDR technologies are an attractive proposition for mobile communication systems. In particular, SDR systems have the potential to offer reconfigurable, multi-mode operation capabilities. A reconfigurable SDR provides the potential to upgrade or enhance the functionality of equipment without the need to change or modify the hardware. In essence, the enhanced functionality is simply ‘downloaded’ onto the equipment. In some cases, this operation may be completely transparent to the user.

We conclude that an SDR is a radio platform whose signal processing path is performed, all or in part, using programmable devices such as (but by no means limited to) FPGAs, digital signal processors (DSPs) and general-purpose processors. As such, the attributes of the transmitted/received waveform can be changed in software and without any physical changes to the hardware. Ultimately, as more and more of the signal path is digitised, SDRs may offer superior tuning capabilities than more traditional radio architectures. However, this is not a mandatory requirement of an SDR. Similarly, an SDR may permit software updates to be applied post-manufacture. Such updates may be downloaded from the Internet, uploaded from plug-in media or, ultimately, transmitted over-the-air. Again, this feature and, if supported, its method of implementation are not mandatory requirements of an SDR.

1.5

Cognitive Radio

The concept of a CR was originally coined by Joseph Mitola III in his doctoral thesis submitted in 2000. Broadly speaking, a CR is a radio that can monitor its local radio environment and/or its geographical position and, using these data, adjust its characteristics in order to optimise its performance. Thus, a CR may, for example, adjust its operating frequency to take advantage of unused spectrum and/or adapt its transmit power, modulation scheme or other waveform parameters to reach an acceptable compromise between quality of service (QoS) and spectral requirements.

It is in this manner that CR has the potential to significantly improve the efficiency with which RF spectrum is used. On the one hand it might search for and use underused spectrum, thereby relieving demand in congested bands. On the other hand it might dynamically optimise its waveform configuration, thereby allowing more intensive reuse of spectrum either through improved resilience to interference received from other spectrum users or reduced interference caused to other users.

Conceptually, therefore, it can be envisaged that developments in SDR will drive developments in CR. As a consequence, this will drive developments in dynamic spectrum allocation (DSA). Therefore, in the long-term, these concepts will give rise to more efficient management of the radio spectrum and include dynamic spectrum trading.

The link between CR and SDR is therefore as follows: CR is the manifestation of the intelligence to manage the generic radio platform, with SDR representing the very flexible, generic radio platform.

1.6

Background

This Main Document is the culmination of three studies [4], [5] and [6], which have been undertaken to categorise and outline the effectiveness of SDR; it includes a discussion on CR.

In today's world, the demand for new radio frequency (RF) spectrum resources is relentless. Historically, licensed RF spectrum in the United Kingdom (UK) has been allocated through a 'command and control' structure, with the regulator, i.e., Ofcom, assigning transmission rights to individual parties. Whilst this approach has many advantages, a significant drawback is that large sections of the RF spectrum are underutilised. In the past this has not been too significant a problem because advances in technology have enabled 'new' spectrum in higher frequency bands to be allocated as demand has grown. However, this trend cannot continue indefinitely.

Following the findings of Professor Martin Cave [7] and, more recently, its spectrum framework review [8], Ofcom is keen to modify its long-term approach to RF spectrum management in order to stimulate a more efficient use of the radio spectrum in the UK. Key to this are the concepts of spectrum trading and spectrum liberalisation. By enabling spectrum users to trade transmission rights and by relaxing certain constraints governing exactly what systems can be operated in a specific band, Ofcom wishes to enable so-called 'market mechanisms' to take over from a command and control approach as the primary method of RF spectrum management in the UK. Ofcom also plans to allocate additional spectrum to licence-exempt short-range applications, e.g., Bluetooth and WiFi. This is shown in Figure 1-3.

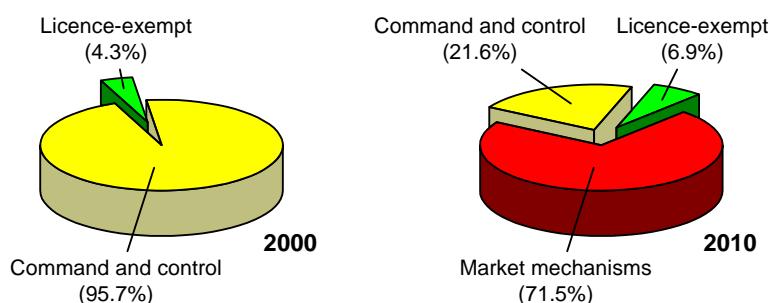


Figure 1-3: Existing (left) and proposed (right) method of RF spectrum management in the UK [8]

Once restrictions controlling what specific frequency bands can be used for are relaxed and spectrum becomes a tradable asset, the incentives are in place to promote more efficient spectrum usage. This holds true for both the primary licence holders, who can sell or lease unused spectrum, and potential leases, who will wish to keep their spectral requirements to a minimum.

UNCLASSIFIED

The proposed changes to the regulatory environment can only lead to the desired improvements in spectral efficiency if flexible radio platforms capable of exploiting a more dynamic radio environment can be developed. Reconfigurable software defined radios (SDRs) have the potential to provide the flexibility necessary. By adding a suitable control entity with limited intelligence to such a radio platform, cognitive radios (CRs) might be realised. A cognitive radio uses knowledge of the local radio environment to intelligently adapt its functionality to best meet the requirements of the user whilst making efficient use of the available radio resources

An example of an SDR development programme often discussed is that of the Joint Tactical Radio System (JTRS), a United States (US) Department of Defence (DoD) initiative. A requirement of the JTRS is operation over the range of 2 MHz to 2 GHz [3], with a possibility of extending to higher frequencies. This requirement is sometimes taken to be a requirement of SDR itself; this is not true. It is true that a radio platform with a ‘highly flexible’, software defined baseband signal path would benefit from a similarly flexible RF front-end. However, the ‘all-singing, all-dancing’ radio platform is just one potential application of SDR. Furthermore, as this report will explain, operation over such frequency ranges will, for the foreseeable future, still require the use of an analogue RF front-end. This said, meeting the JTRS requirement is an interesting and challenging problem. Therefore, where appropriate, some sections of this report will consider the challenges that such a requirement might raise.

1.7 Consortium Members

This study was completed through the collaboration of a consortium consisting of five members. Leading this consortium was QinetiQ. The consortium members are introduced below:



QinetiQ

QinetiQ has been actively researching software defined radio since its inception in 1996 and is a member of the Software Defined Radio Forum (SDRF) which gives excellent visibility of the commercial implementation of these technologies. QinetiQ’s contribution to this report includes antennas, software, security, radio management, including cognitive radio and an assessment of application and areas of deployment for SDR. QinetiQ is responsible for the overall management of this study as well as for the compilation and editing of the report.

UNCLASSIFIED



Multiple Access Communications (MAC) Ltd

Founded in 1986, MAC Ltd is an independent private company founded to provide the telecommunications industry with high-quality consultancy, and a concomitant range of products. The company operates internationally, particularly in the USA, Europe and the Far East and its customers include network operators, equipment manufacturers, government departments and semiconductor houses. MAC Ltd is a member of ETSI and TETRA Memorandum of Understanding (MoU). MAC Ltd's contribution to this report includes the regulatory issues, commercial driver studies and an assessment of spectral efficiency gains for SDR and CR.



University of Southampton

The first of three academic consortium members is the Department of Electronics and Computer Science at the University of Southampton. The University of Southampton's contribution to this study includes the antenna processing and waveform studies.



The University of York

The second academic consortium member is the Department of Electronics at the University of York. The University of York's contribution to this study includes the multiple-input, multiple-output (MIMO) study.



Vienna University of Technology

The final consortium member is the Vienna University of Technology. The Vienna University of Technology's contribution to this study includes the RF linearisation study.

1.8 Report Structure

The report structure is outlined as follows:

- Chapter 1: Introduction

This chapter.

- Chapter 2: Antennas

We start in Chapter 2 by considering antennas. Here, we look at the requirements of the antenna in wideband SDR applications. This section also includes a brief overview of antenna fundamentals to allow the subsequent discussions to be analysed. We also describe various antenna solutions that might be applicable to wideband SDR applications, both now and in the future.

- Chapter 3: RF Linearisation

Most ‘modern’ modulation methods require the use of ‘linear’ power amplifiers. However, linear power amplifiers are typically less efficient than nonlinear designs, but when linear power amplifiers are driven to saturation they exhibit non-linearities. In Chapter 3, we discuss modern RF ‘linearisation’ methods and present some simulation data to illustrate the advantages of using such techniques.

- Chapter 4: Antenna Processing

The performance of the signal path between the analogue RF signal at the antenna and the digital domain is critical in an SDR on both the receive and transmit signal paths. Chapter 4 considers antenna processing. Specifically, this section describes potential SDR receiver and transmitter architectures and discusses the performance of current ‘state-of-the-art’ analogue-to-digital and digital-to-analogue conversion devices. Finally, this section introduces some novel concepts that have the potential to dramatically enhance the performance of these conversion devices.

- Chapter 5: MIMO Technology and SDR

Chapter 5 discusses the concept of MIMO. MIMO attempts to exploit a multi-path environment to enhance channel capacity. Although not a requirement of SDR, it might be argued that MIMO is a complementary technology to SDR. Furthermore, the potential flexibility of SDR means that SDR might well prove to be an enabling technology for MIMO. As well as an in-depth discussion of MIMO, this section presents example simulation data to illustrate the potential gains to be found by employing MIMO techniques.

- Chapter 6: Waveforms

The promise of achieving interoperability between communication devices that are nominally targeted at different radio technologies has driven interest in SDR technology. Chapter 6 focuses on waveforms and their potential for integration. In this section we analyse key waveform characteristics that might impact on the ability to integrate them. We also give a brief overview of waveforms as defined in some of today’s major wireless standards, and review modulation methods and the topic of waveform integration.

UNCLASSIFIED

- Chapter 7: Software Aspects

In Chapter 7 we move to consider software aspects associated with SDR. Here we discuss in some depth existing standard software architectures and models and evaluate their relevance and application to SDR.

- Chapter 8: Security

Chapter 8 considers software security aspects associated with SDR. In particular, this section considers the risks that might be associated with over-the-air software updates and discusses approaches and methods that might be employed to help protect such data transfers.

- Chapter 9: Radio Management and Cognitive Radio

Chapter 9 discusses the topics of radio management and cognitive radio (CR). In this section we discuss CR and how it might be used to improve the effective use of the available radio spectrum. As a basis for these discussions we also consider the topic of radio management. Radio management is about optimising the use of the available radio spectrum without causing unacceptable inter-technology interference to others.

- Chapter 10: Regulatory Issues

Chapter 10 discusses regulatory issues associated with the development and deployment of SDR. We start by reviewing SDR with a view to defining what needs to be regulated in an SDR. We also briefly consider why SDR is an important step in the evolution of radio and why regulatory hurdles should not be allowed to deter the continued development and adoption of SDR. The regulatory environment in both the US and Europe is reviewed, giving particular attention to the Federal Communications Commission's (FCC's) efforts to acknowledge and adapt to the emergence of SDR. Finally, in this section we briefly discuss the concepts of spectrum trading and liberalisation and how SDR might help exploit their full potential.

- Chapter 11: Commercial Drivers

Although SDR receives considerable academic interest, it is important to determine that SDR is not simply a purely academic vision and interesting research subject. Chapter 11 considers the commercial drivers behind SDR. In this section, we discuss the potential advantages (and disadvantages) of SDR from the point of view of the manufacturer, network operator and end user. The discussion made in this section draws on industry feedback gathered during a 'Stakeholders' meeting held in January 2005.

- Chapter 12: Assessment of SDR's First Applications and Areas of Deployment

The following areas have been identified as potential markets for SDR and are examined in Chapter 12: Cellular networks -Local and metropolitan area wireless networks; Transportation - Automotive and Commercial; and Public Safety. With current technology there are limitations on how wide a frequency range a radio device can operate over and there are also restrictions on how much of the radio's functionality can be implemented in software. The progress of SDR is explored in some detail.

UNCLASSIFIED

- Chapter 13: Spectrum Efficiency Gains of SDR and CR

In Chapter 13 we consider the spectral efficiency gains of SDR and CR. The potential of SDR to improve spectral efficiency is considered in three areas. First we consider cognitive, multi-mode terminals. Second we consider dynamic spectrum allocation and finally we consider the use of dynamic waveforms to help maximise the spectral efficiency of a particular system. Whilst we have considered these concepts in isolation, we note that they might be combined in practice to great effect. Thus, for example, we might envisage a multi-mode terminal implementing new, spectrally efficient, dynamic waveforms incorporating dynamic spectrum allocation as well as various legacy waveforms.

- Chapter 14: Conclusions

Finally, in Chapter 14, the main findings of this study, including epoch estimations for the development of SDR, are summarised.

1.9 Reference

As stated in Section 1.1, this Main Document is the detailed version of the concise Overview Document [1].

References can be found in both the References/Bibliography Section and in Appendix B.

1.10 Acknowledgements

The following people are acknowledged for their contributions and suggestions in this study: Ahmad Atefi, Daniel Bradford, Neil Briscombe, Alister Burr, Julie Fitzpatrick, Nick Frall, Peter Gould, Tim James, Markus Rupp, William Webb, Stephan Weiss and Gavin Wood.

2 Antennas

By Gavin Wood, QinetiQ.



2.1 Introduction

In this chapter we consider antennas, an essential component of any radio system. In particular, we will look at advanced antenna systems that are most likely to be applicable to software defined radio (SDR) systems.

We begin by discussing the key components of an SDR system before focusing on the antenna component itself. We will follow this with a brief introduction to some fundamental aspects of antenna operation upon which further discussions will be made. Subsequently, we will identify and discuss a range of antenna solutions that might be suitable for use in SDR systems. Finally, a forecast is provided, indicating the epochs for the developments of antenna technology.

2.2 Software Defined Radio

Looking towards the future, SDR technologies are an attractive proposition for mobile communication systems. In particular, SDR systems have the potential to offer reconfigurable, multi-mode operation capabilities. A reconfigurable SDR provides the potential to upgrade or enhance the functionality of equipment without the need to change or modify the hardware. In essence, the enhanced functionality is simply ‘downloaded’ onto the equipment. In some cases, this operation may be completely transparent to the user. For the purposes of this section, we will assume the following definitions for the terms ‘software defined radio’ and ‘radio’.

Software defined radios are “radios that provide software control of a variety of modulation techniques, wideband or narrowband operation, communications security functions (such as frequency hopping), and waveform requirements of current and evolving standards over a broad frequency range”.

The term ‘radio’ is taken to mean “any system that seeks to communicate with another system by means of a modulated signal within the RF spectrum”. Typical examples might be a mobile phone, a personal digital assistant (PDA), a mobile radio or a wireless-enabled laptop computer.

The communications industry is seeking to create radios that can handle multiple frequency bands, use multiple transmission protocols and be reconfigurable, preferably on-the-fly. These systems provide three key advantages over traditional ‘hard-wired’ radios in that they:

- provide low-cost solutions since functionality that used to be hard-wired into the radio can now be provided via software
- are easy to upgrade either by a physical connection or even remotely, i.e., over-the-air, by radio transmission
- potentially allow a faster evolution of industry standards, end-user equipment and communication and network infrastructures.

UNCLASSIFIED

In addition, there are issues arising from the considerable strain on existing bandwidth and infrastructures from the growing number of mobile and wireless devices. With ever increasing numbers of users requiring access to different parts of the spectrum, it is hard to envisage that any of these issues can be solved with conventional, inflexible hard-wired radio systems.

2.2.1 Key Components

The three main components of a simple SDR are shown in Figure 2-1. These are:

- transmitting and receiving antenna (a single antenna may perform both functions)
- transmit and receive amplification and frequency translation
- digitisation and digital signal processing (DSP).

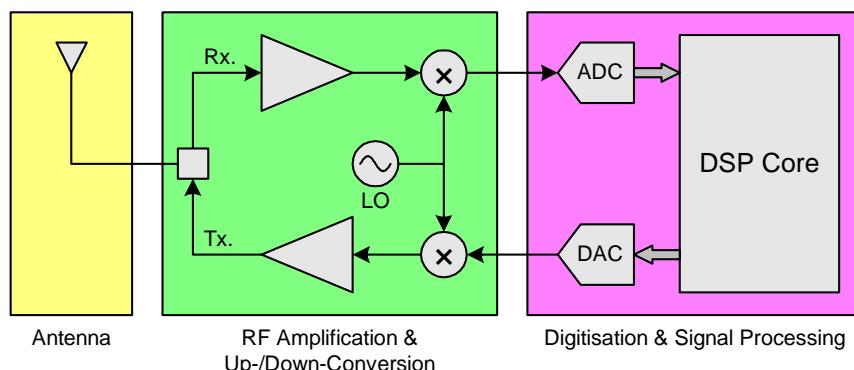


Figure 2-1: Main components of a simple SDR

It is the staggering rate at which the speed of processors and DSP components has increased that has led to the realisation of SDR concepts. The reconfigurable DSP core allows the processing functions to be easily changed on-the-fly to incorporate different functionality as required.

Conversely, it is not so straightforward to integrate the same level of reconfigurability into the radio frequency (RF) amplification and the up and down conversion stages. However, recent advances in analogue-to-digital and digital-to-analogue converters (ADCs and DACs, respectively) have made it possible to directly convert signals closer to the antenna, at high speed and with increasing dynamic range. This is a huge step towards a completely digital solution and consequently optimal flexibility. A number of issues associated with the main components of SDRs are discussed briefly below.

UNCLASSIFIED

2.2.1.1 Digitisation and Signal Processing Technology

The reconfigurability of the radio system is determined by the flexibility of the software and the complexity of the DSP core. DSP technology is progressing very rapidly (Section 4.6 provides a discussion on the rate of progress, and a forecast, for this technology). The primary issue at present is that of power consumption. ADC technology is also progressing rapidly, but again power consumption is a key issue. A radio system covering all personal mobile radio and mobile phone frequencies would need to operate between 100 and 2200 MHz [9]. This bandwidth is increased to include frequencies down to 20 MHz and below when considering the requirements of the military Joint Tactical Radio System (JTRS). This is a large input bandwidth for current ADC components, but a suitable baseband down-conversion scheme prior to digitisation would relax this requirement.

2.2.1.2 RF Amplification and Conversion Technology

The main components of interest in this are the duplexer or circulator and RF power amplifiers. The duplexer may take the form of a simple switch, which has advantages including wide bandwidth. However, there may be issues with isolation, and the transmitter may need to be disabled to avoid saturating the receiver in receive mode. If this is the case, the transceiver cannot transmit and receive simultaneously which would be a significant limitation. The use of a circulator would allow simultaneous transmit and receive operation, but these devices generally have narrow bandwidth and poor isolation. They are generally used to isolate the antenna from the transmitter to relax the specification on antenna input match. Considerable research has been conducted into the linearisation of power amplifiers and a number of techniques exist and are widely employed. For narrowband systems, typical intermodulation product levels of -70 dBc can be achieved with conventional techniques. If this is to be extended across a broader band, more sophisticated pre-distortion and feedforward techniques need to be applied [9]. Numerous receiver architectures may be employed to deal with these various issues [10].

2.2.1.3 Antenna Technology

It is alarming, but perhaps understandable, that antenna technologies for SDR applications have not been widely examined. They do not generally appear as a separate block in the ‘system design’ as most RF systems are relatively narrow band. RF system designers are broadly more concerned with antenna gain/directivity, radiating characteristics, input match characteristics (or voltage standing wave ratio (VSWR)) and cross-polar rejection. Of course, this is a sweeping generalisation, but many applications require bandwidths that are readily achievable using existing antenna designs.

The ratio between the upper and lower passband edges of current ‘wideband’ antenna designs ranges typically between 15:1 and 25:1 (less than five octaves). For the example specification given above, consider a bandwidth of, say, 20 to 2000 MHz, a frequency range ratio of 100:1 is required (nearly seven octaves). This will almost certainly prove to be a difficult specification to meet with a single antenna.

UNCLASSIFIED

2.2.2

SDR Antenna Requirements

There is a significant thrust towards SDR solutions in the military arena, more specifically the JTRS concept, which aspires to achieve universal wireless military communications integrated within an SDR regime. The bandwidth likely to be occupied by such a system is 2 to 2000 MHz (with the possibility of extending to higher frequencies), covering the majority of the existing RF bands currently in use. However, there is a growing interest in both military and civil markets in SDR applications operating above 2 GHz, potentially extending up to 50 GHz. Whilst higher operating frequencies bring about increased path losses, system complexity and expense, an ever increasing number of users requiring bandwidth will ultimately demand the use of a wider portion of the electromagnetic spectrum [11]. Furthermore, the growth in demand of data services is pushing current systems to capacity.

2.3

Antenna Fundamentals

In order to fully appreciate the underlying issues with wideband antenna design, it is necessary to first understand some important, fundamental antenna characteristics.

The fundamental characteristics of an antenna are its gain and half power (i.e., -3 dB) beamwidth. The theorem of reciprocity dictates that the transmitting and receiving properties of an antenna are identical for a given frequency.

The gain is a measure of how much power at the input of an antenna is radiated in a particular direction. Antenna gain can be expressed either with respect to a theoretical isotropic radiator (gain expressed in dBi) or the maximum gain of a theoretical dipole (gain expressed in dBd). An isotropic radiator is a conceptual antenna element that radiates equally in all directions. 0 dBd is equivalent to 2.15 dBi. So for a given direction vector in space (θ, ϕ) , the gain is given by:

$$G(\theta, \phi) = 10 \times \log_{10} \left(\frac{dP/d\Omega}{P_{in}/(4 \cdot \pi)} \right) \text{ dBi} \quad \text{Equation 2-1}$$

where P_{in} is the total input power and dP is the increment of radiated power in solid angle $d\Omega$. The input power is given by:

$$P_{in} = \frac{E_a^2 \cdot A}{\eta \cdot Z_0} \text{ W} \quad \text{Equation 2-2}$$

where E_a is the average electric field over the area A of the antenna aperture, Z_0 is the impedance of free space ($120\pi \Omega$) and η is the net antenna efficiency. The output power dP over a solid angle $d\Omega$ is:

$$dP = \frac{E^2 \cdot r^2}{Z_0} \cdot d\Omega \text{ W} \quad \text{Equation 2-3}$$

UNCLASSIFIED

where E is the electric field at a distance r . By diffraction theory, the electric field E along the boresight direction is:

$$E = \frac{E_a \cdot A}{r \cdot \lambda} \text{ V/m} \quad \text{Equation 2-4}$$

where λ is the wavelength. The boresight gain can now be determined in terms of the physical size of the antenna by combining Equation 2-1, Equation 2-2, Equation 2-3 and Equation 2-4:

$$G = 10 \times \log_{10} \left(4 \cdot \pi \cdot \eta \cdot \frac{A}{\lambda^2} \right) \text{ dBi} \quad \text{Equation 2-5}$$

The net efficiency η depends on the electric field distribution over the antenna aperture, and the total radiation efficiency associated with various losses, e.g., ohmic loss, phase non-uniformity, surface roughness and cross-polar effects. For a typical antenna $\eta = 0.55$.

Equation 2-5 highlights some important points. Firstly, gain is proportional to area, so for a given frequency, a larger antenna has more gain and therefore radiates more power over a smaller solid angle. Secondly, the gain is inversely proportional to the square of wavelength, so for a given antenna aperture, the gain increases as frequency increases.

The half power beamwidth is the angular separation between the half power points on the antenna radiation pattern, where the gain is half (i.e., -3 dB) the maximum value. For a reflector antenna, the half power beamwidth may be expressed in radians as:

$$BW_{-3dB} = \frac{k \cdot \lambda}{L} \text{ rad} \quad \text{Equation 2-6}$$

where L is the length of the aperture in the plane being measured and k is a factor that depends on the shape of the reflector and the illumination method used. Typically, however, a value of 1.22 is often used for k .

Figure 2-2 shows a graph of two antenna patterns as a function of angle. The antennas both have a 1 m circular aperture, with the first antenna operating at 1 GHz and the second at 4 GHz. It can be seen that the -3 dB beamwidth of the 1 GHz antenna is four times that of the 4 GHz antenna.

UNCLASSIFIED

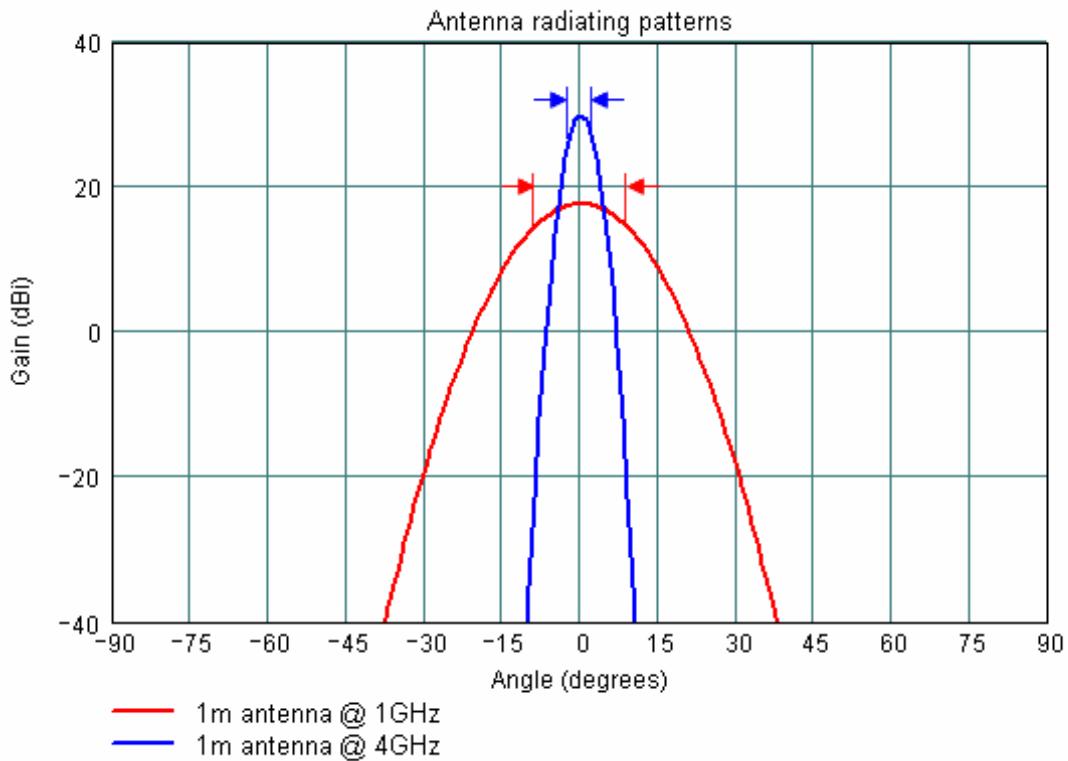


Figure 2-2: Antenna radiating patterns for 1 m antennas operating at 1 GHz and 4 GHz

Figure 2-3 shows the same information but plotted in polar form. In both figures, it is clear that the 4 GHz antenna has 12 dB more gain than the 1 GHz antenna. So if an antenna application requires more gain, the designer has the option to either increase the physical size of the antenna, or increase the operating frequency.

UNCLASSIFIED

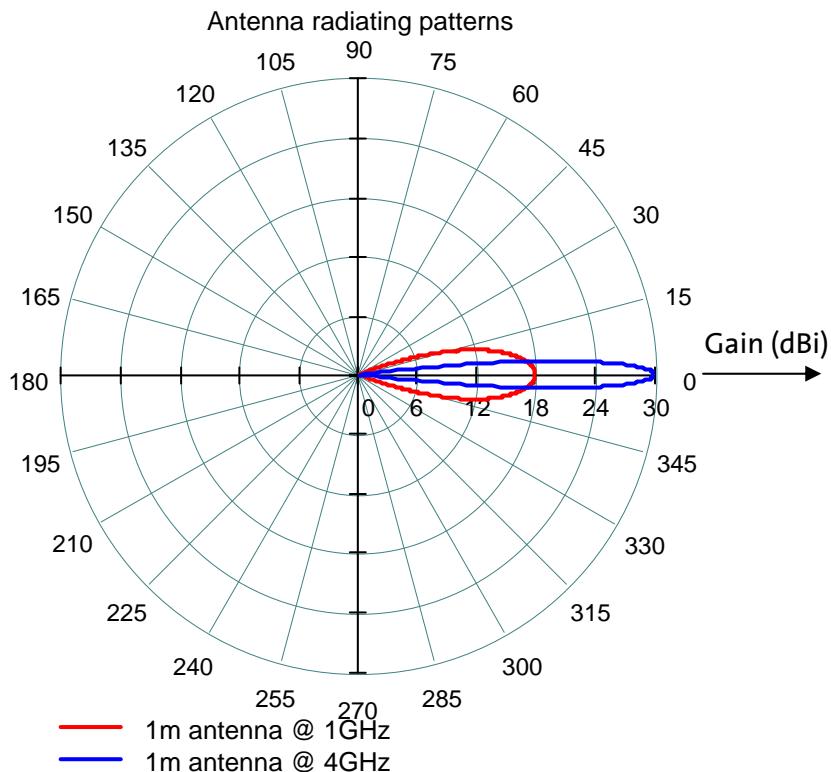


Figure 2-3: Polar plot of antenna radiating patterns for 1 m antennas operating at 1 GHz and 4 GHz

The ramifications of these fundamental characteristics are that even if a single antenna *could* be designed to provide a good match over the JTRS frequency range of 2 to 2000 MHz, the antenna would have non uniform characteristics over its operating bandwidth. At the lower end of the band the antenna would have low gain and be almost omni-directional (i.e., have a very wide half power beamwidth). At the higher end of the band, however, the antenna would have high gain and a narrow beamwidth. This is an important consideration when designing an antenna to cover a wide bandwidth and a number of different applications with potentially differing antenna requirements.

2.4 Antenna Options for SDR

In this section we will describe the types of antennas and antenna systems that are most likely to be applicable to wideband SDR systems. We will consider both mobile handset and base station applications, outlining the advantages of the various antenna options. We will also highlight some of the areas in which antenna research is currently focussed.

UNCLASSIFIED

2.4.1 Phased Array Antennas

The ability to distinguish between, and separate, users in a communications system is essential. The most common multiple access schemes currently employed are frequency-division multiple access (FDMA), time-division multiple access (TDMA) and code-division multiple access (CDMA). These schemes separate users into the frequency, time and code domains, respectively, giving three distinct degrees of diversity.

A smart antenna system seeks to add an additional degree of diversity to reduce interference between users and consequently increase user capacity through dynamic adaptation of the antenna's radiating properties [12], [13]. The most obvious example of this is a phased array antenna, an array of separate radiating 'elements' whose signals, when added together, form a beam. Flexibility and control over the beam shape is obtained during the beamforming process by altering the amplitude and phase of the individual element responses prior to summation.

Phased arrays provide spatial diversity in that the main beam can be steered in a chosen direction whilst 'nulls' can be steered in other directions simultaneously. Thus, the antenna can be adapted to give high sensitivity to the signals received from one user whilst suppressing those from other users. This can be used to suppress interference between other SDR units and to reject sources of jamming and other hostile interference in military systems. Note that, although we have here considered the receiving of signals, the same principles can also be applied to shape the radiation pattern of the transmitted signal. Two schemes utilising the beam shape flexibility are generally adopted, these are beam-steering and beam-switching.

2.4.1.1 Beam-Steering Systems

In a beam-steering system, the main beam gain can be steered dynamically to a chosen direction by altering the 'phase front' applied to the elements in the array during the beamforming process. Thus, the main beam gain from a base station can be steered in the direction of an individual mobile user. As the location of the mobile user changes, the beam steer can be updated to keep the user 'spotlighted'. Figure 2-4 shows an antenna with three beams steered to different directions.

UNCLASSIFIED

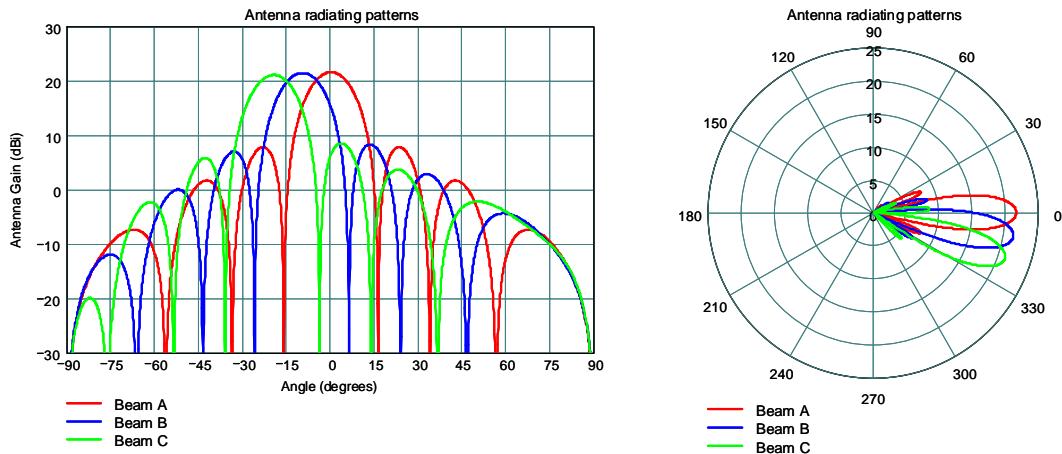


Figure 2-4: Antenna beams steered to different directions

Note that, for the different steer angles, although the beam shapes are similar they are not identical. This is due to the fact that the projected array aperture decreases with increasing scan angle. As the array aperture decreases, the beamwidth widens in accordance with Equation 2-6. This is shown in Figure 2-5.

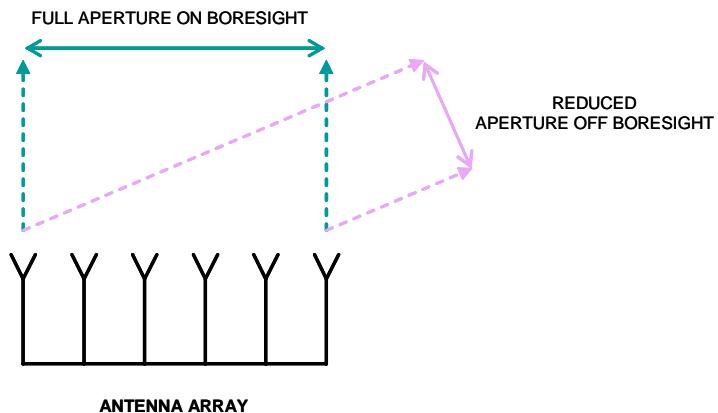


Figure 2-5: Reduction in aperture with beam steer angle

Beam-steering offers significantly higher interference reduction performance over beam-switched systems (discussed in the next section), but are inevitably more complex and expensive to implement. Sources of interference can be identified and nulls in the beams steered to those directions [14]. Furthermore, because the beams can ‘pan’ to follow a particular subscriber, the full main beam gain is directed at the subscriber at all times. This potentially allows the mobile unit to have lower gain and conserve power. In addition, special algorithms can be employed to resolve separate multi-path signals and recombine them to optimise signal-to-interference-plus-noise ratio (SINR).

UNCLASSIFIED

2.4.1.2 Beam-Switching Systems

Beam-switching uses a number of fixed beams steered to predetermined directions. A mobile subscriber switches between the beams according to its location. In its simplest implementation, only a basic switching function between the different predefined beams is required. The switching mechanism should quickly and transparently switch a mobile subscriber to the correct beam. This ‘handover’ is usually accomplished by receiving the subscriber’s signal on multiple beams and choosing the best signal based on a measurement of the signal-to-noise ratio (SNR). Unlike the beam-steering method, no subscriber tracking is required or employed.

An advantage of this type of system is that the predetermined beams can be set up in such a way as to minimise interference to other beams. Consider Figure 2-6 below. The three beams from the antenna array can be set up such that any two beams have a null along the direction of the third beam main gain direction.

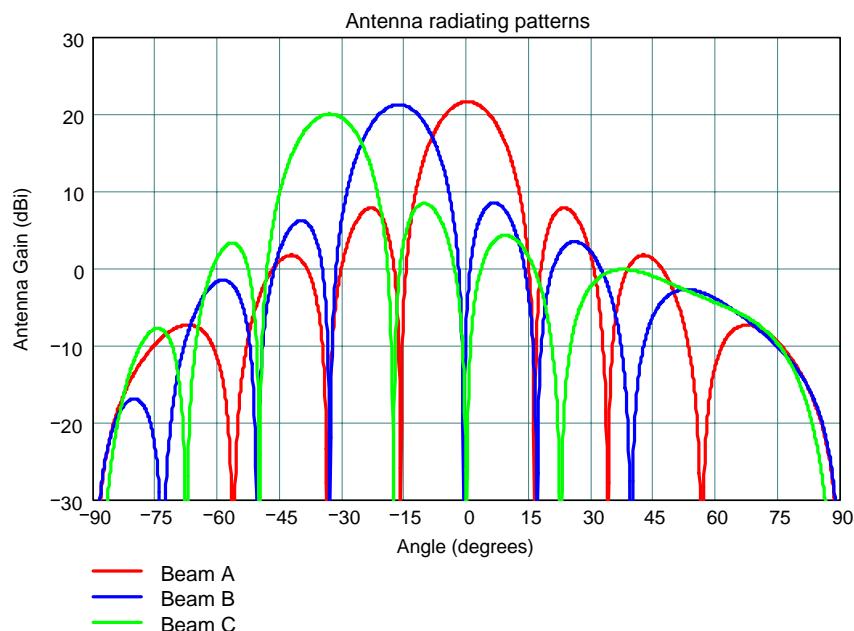


Figure 2-6: Three antenna beams with nulls along main beam direction

There are limitations to beam-switched systems. The beams are predetermined and do not change properties dynamically, so the signal strength varies as a subscriber moves through a beam; towards the edge of a beam, the signal strength can roll off rapidly before the subscriber is switched to the next beam. Also, it is possible, under certain circumstances, for interference from subscribers in a higher gain region of a beam to swamp the signals of others in lower gain regions of the same beam. The sidelobe levels of the beams are critical in obtaining minimal interference between beams.

UNCLASSIFIED

Both beam-steering and beam-switching are currently used in universal mobile telecommunications system (UMTS) mobile telephone base stations. The base station masts are triangular in shape and have an antenna array on each side. Each antenna covers a sector of 120°. Mobile units move in and out of different sectors and are effectively switched between beams as described above. Beam-steering has recently been implemented by some operators. Steering the beams vertically (i.e., electrically adjusting the 'down-tilt') alters the size and shape of the cell, i.e., the region on the ground illuminated by the beam, so a dynamic rather than a static cell structure can be achieved with fixed base stations. This has significant advantages in network planning, where positioning base stations optimally can prove difficult. Also, the cell structure can change temporally as mobile user density migrates from cell to cell, e.g., a city centre heavily populated with business and retail premises generally has a high user density during the day, but a low user density at night.

Phased array technology is most likely to be of benefit to base station type applications, but may also be beneficial if employed to a limited degree in handsets. An antenna that can dynamically steer maximum gain to a chosen direction and minimise power in others will be attractive to covert military communication systems. Advanced signal processing techniques available to an array antenna, but not to conventional schemes, provides potential for multi-path mitigation, reduced interference and increased system capacity.

2.4.2

Reconfigurable Antennas

Sections 2.2.2 and 2.3 have highlighted the wide bandwidth likely to be required by SDR applications. Covering this bandwidth with a single antenna that has sufficient efficiency is a considerable challenge. Since the antenna gain is fundamentally linked to the physical size of the antenna (see, Equation 2-5), so too is the gain-bandwidth product [15]. To exacerbate the problem, SDR handsets are likely to require wideband antennas in physically small footprints, such as those available in mobile telephones.

One possible solution to this problem is to use reconfigurable antennas that can be tuned to different frequency bands, whilst maintaining sufficient instantaneous bandwidth and efficiency within each band. Antennas in this class do not cover all bands simultaneously, but provide dynamically selectable narrower instantaneous bandwidths at higher efficiencies than conventional antenna designs.

One type of reconfigurable antenna is the shorted patch antenna reported in [16]. This antenna occupies an area of approximately 30mm square and is tuneable across a number of bands from 800 MHz to 2000 MHz, with input matches in each band of better than -10 dB. The antenna is approximately a quarter of a wavelength at the fundamental resonant frequency and the structure is as shown in Figure 2-7.

UNCLASSIFIED

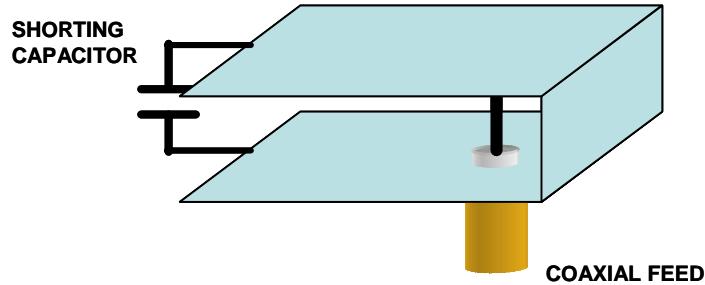


Figure 2-7: Shorted patch antenna

Adding a capacitor across the main radiating edge reduces the resonant frequency. It is the value of this capacitance that determines the radiating frequency. A limitation of this type of antenna is the Q-factor of the capacitor, which determines the instantaneous bandwidth at the tuned frequency.

Another class of reconfigurable antenna involves mechanically tuning the antenna by changing its shape rather than electrically tuning it as in the shorted patch antenna. Consider the simple dipole in Figure 2-8. The frequency of operation will depend on the length of the 'arms' of the dipole. These arms can be made longer or shorter by switching in or out the extra lengths, A, B and C. The electronic switches could be relays, or in the case of small antennas, manufactured using micro-electromechanical system (MEMS) technology [17].

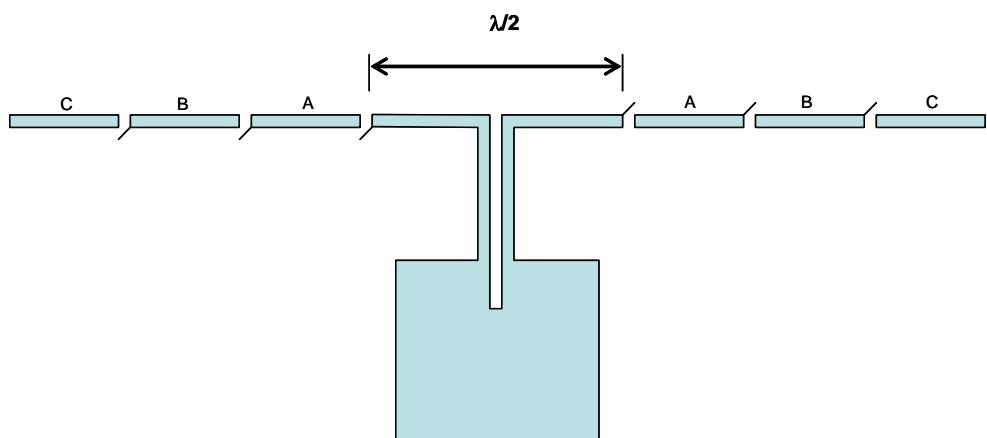


Figure 2-8: Dipole antenna with switchable arm lengths

Fathy et al [17] have taken this a step further by actively defining an antenna aperture using biased P-i-N diodes. An array of P-i-N diodes is manufactured on a silicon substrate in a lattice fashion, although this need not be regular. Activation of the diodes creates a conductive region in the silicon around the diode, thus the size and shape of the conductive region defines the antenna aperture. Figure 2-9 shows two examples of how the aperture may be defined electronically, as patch or bow-tie type antennas.

UNCLASSIFIED

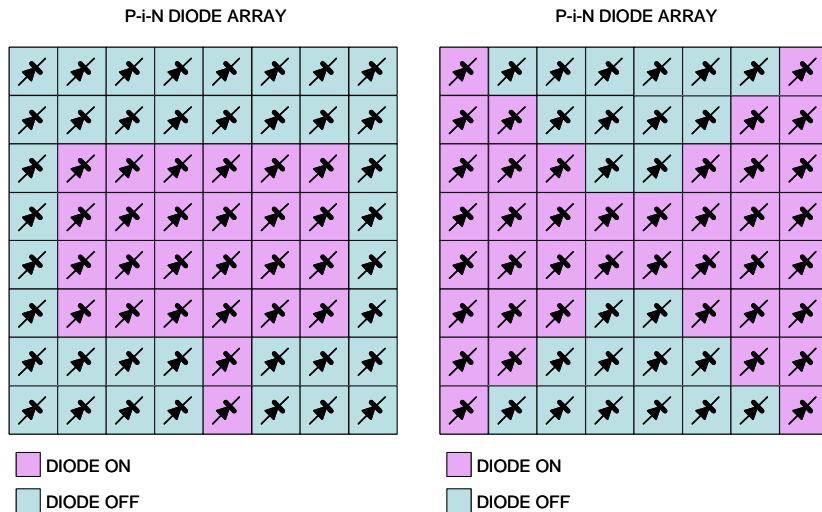


Figure 2-9: Array of switchable P-i-N diodes allows reconfigurable aperture shape

A current limitation with this type of antenna is that the six inch diameter aperture of the silicon processing wafers in a standard foundry run imposes a minimum operating frequency of 2 GHz. Also, the complexity of the diode switching circuitry may render this antenna unsuitable for some applications. However, this type of technology presents a significant advancement towards a truly reconfigurable antenna.

An antenna that can be reconfigured on-the-fly, such as those mentioned above, allow some interesting possibilities for SDR applications. The size and shape of the antenna aperture can be altered to match the frequency and gain requirements of different applications, especially if these requirements are known in advance so that the appropriate commands can be sent to the diode or relay switching circuitry. In addition, it has been proposed by Linden [18] that if a suitable control loop can be implemented, the antenna characteristics might be updated dynamically to optimise receiver SNR. Figure 2-10 shows a control loop that uses a genetic algorithm to evolve the antenna shape. The algorithm uses the signal strength from the receiver circuit and alters the shape of the antenna until an optimal SNR is obtained.

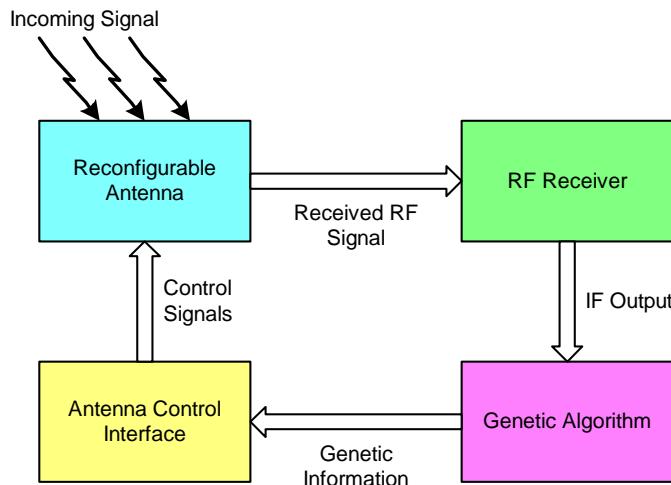


Figure 2-10: Evolvable antenna control loop

A possible limitation is the time taken for the antenna to ‘evolve’ into the optimal shape. It takes approximately 100 cycles of the control loop to achieve a stable antenna shape. Therefore, the control loop needs to be fast if the antenna is to adapt and evolve in a dynamic environment. As with any control loop, there is a danger of instability. Alternatively, the number of evolution cycles could be limited to improve the settling time, but at the expense of an optimal shape.

Reconfigurable antennas are an exciting but relatively immature technology. Consequently they are not currently employed in practical radio systems. However, current research is focussed on developing this technology for use in cars as an adaptive antenna to receive traffic information, global positioning system (GPS) updates and possibly Internet data. Ultimately, they are likely to find use in SDRs where the antenna characteristics cannot be compromised for different applications, e.g., a single antenna operating at 2.4 GHz and 5.4 GHz will have half the beamwidth at the higher frequency. If this is not acceptable, a reconfigurable antenna could be employed to keep the beamwidth constant.

2.4.3 Antennas for Polarisation Diversity

An important feature of antennas that is often overlooked is the polarisation selectivity. Most antennas receive signals on a single polarisation whilst rejecting signals on others. This is exploited in a number of areas, for example, two television transmitter stations in close proximity will transmit on orthogonal linear polarisations to allow receivers to select one transmitter whilst rejecting the other. Similarly, the uplink and downlink information from base stations to mobile subscribers can be sent on different circular polarisations.

It would therefore be beneficial, in some applications, to have the capability of selecting the polarisation on which to transmit and receive signals from an SDR unit. There are antennas capable of performing this function, but at the expense of extra complexity in the feed arrangement.

UNCLASSIFIED

If we take a simple metallic patch antenna with microstrip feed, the patch will radiate on a linear polarisation orthogonal to the feed point. However, if we feed the patch at two points on orthogonal sides of the patch, we can arrange that the patch radiates on a polarisation determined by the phase difference between the feed points. Examples of single and dual feed patch antennas are shown in Figure 2-11.

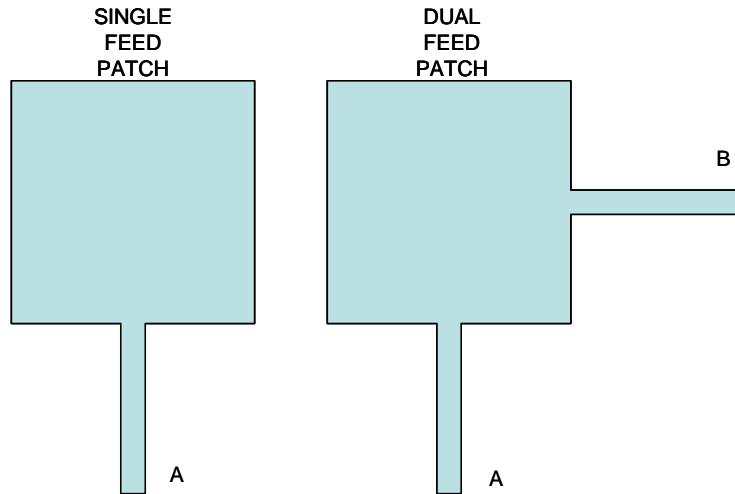


Figure 2-11: Square patch antennas fed under single and dual feed arrangements

Linear horizontal or vertical polarisation can be achieved from the dual feed patch by exciting either A or B ports, respectively. However, if both A and B ports are excited simultaneously, the phase difference between the ports determines the radiated polarisation. Figure 2-12 shows the transmitted polarisation if both A and B ports are excited with equal amplitudes and a phase difference as denoted in the table. Exciting both A and B ports with the same or opposite phase produces linear polarisation at $\pm 45^\circ$. A 90° phase difference produces either left or right-hand circular polarisation.

		FEED A			
		0°	90°	180°	270°
		0°			
FEED B		90°			
180°					
270°					

Figure 2-12: Transmitted polarisation from an antenna excited at two orthogonal ports

UNCLASSIFIED

Linear polarisation at any angle can be achieved by feeding ports A and B in-phase but at different amplitudes. Elliptical polarisations can be achieved through a combination of phase and amplitude differences in the excitations at the A and B ports. These polarisations can be achieved in both transmit and receive modes.

Another antenna in this class is the spiral antenna, shown in Figure 2-13. A spiral antenna is essentially constructed from one or two dipoles that are coiled to form a spiral. This winding of the antenna allows the maximum operating wavelength to be equal to the circumference of the outer ring of the spiral and the antenna to occupy a significantly smaller footprint. However, the structure of the spiral has many resonances making the antenna broadband. The four feed spiral (shown in Figure 2-13 (right)) antenna also allows the full polarisation diversity offered by the dual fed patch.

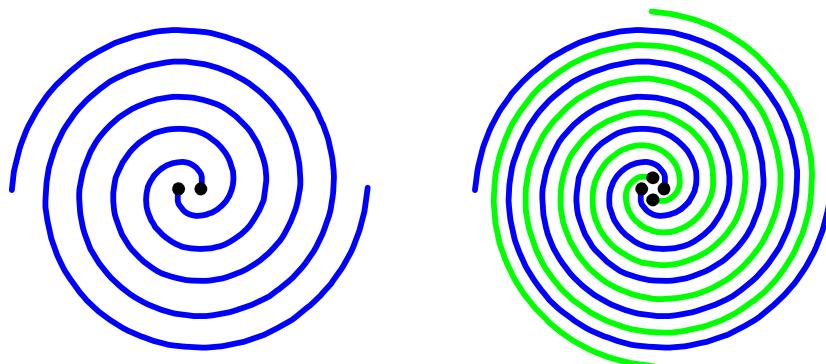


Figure 2-13: Two-arm and four-arm spiral antennas

The sinuous spiral antenna is similar in its operation to the spiral. Both antenna types are detailed in [19]. An example of a sinuous antenna is shown in Figure 2-14.



Figure 2-14: Four-arm sinuous antenna

UNCLASSIFIED

The sinuous antenna can radiate in three additional modes producing different radiating characteristics. Mode 1 is characterised by a phase shift of 90° between adjacent arms. Mode 2 has a phase shift of 180° . Example radiation patterns for Modes 1 and 2 for a 10 cm radius sinuous antenna at 2 GHz, are shown in Figure 2-15. Both modes are circularly polarised in the direction of the main beam.

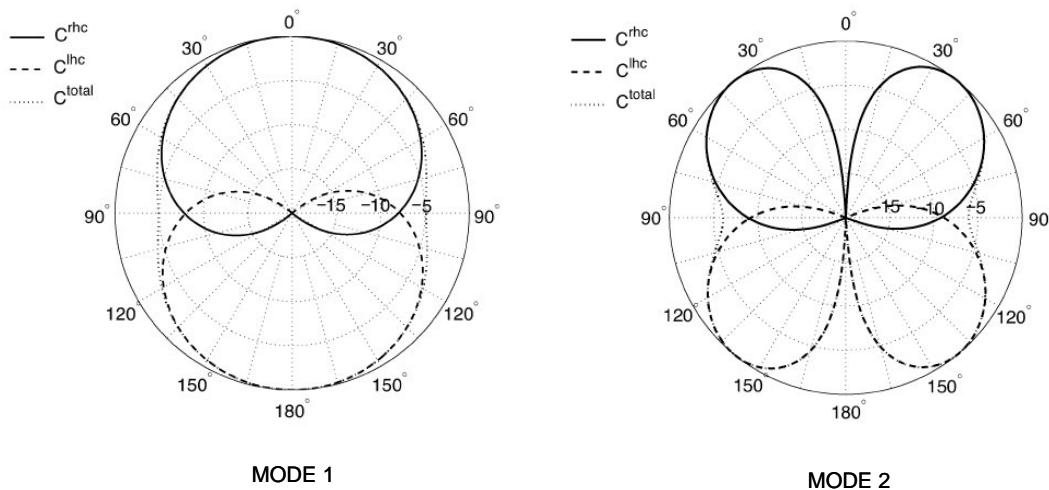


Figure 2-15: Radiation patterns for Modes 1 & 2 of a 10 cm radius sinuous antenna at 2GHz, separated into left- and right-hand circular polarisation [19]

The antenna radiates from an ‘active zone’ located around the centre of the antenna. The size of this zone is determined by the ratio of its circumference to the wavelength. For Mode 1, the circumference of the active zone is one wavelength, for Mode 2 it is two wavelengths. Above the lower frequency limit (wavelength given by the outer circumference of the antenna for Mode 1 and half the circumference for Mode 2), the radiation patterns detailed in Figure 2-15 do not change significantly. When operated in Mode 3 (270° phase shift between adjacent arms), the radiation pattern has the same amplitude response as Mode 1, but with the left and right-hand polarisations reversed, so Mode 1 and Mode 3 are orthogonally polarised.

Antennas in this class exploit multi-mode diversity, i.e., a combination of pattern (spatial selectivity) and polarisation diversity. They are very broadband (bandwidths of multiple octaves are achievable) and highly applicable to multi-standard radios and SDRs. The physical footprint of these antennas is smaller than that of other types, saving space, weight and cost.

Polarisation diversity has a considerable benefit in high multi-path environments such as dense urban terrain. Signals between units become decorrelated in the presence of multi-path. This is usually dealt with in the spatial domain by having several antennas, located many wavelengths apart, receiving the signals and selecting the best signal or, optimally, combining the signals with a diversity combiner. However, multi-path also varies the polarisation of a signal and it is possible to improve performance by receiving on two orthogonal polarisations simultaneously. The advantage is that this can be done with a single antenna.

2.4.4 Multi-Band and Ultra-Wideband Antennas

The vast majority of antennas are designed to operate at a specific frequency and bandwidth. This is usually achieved by having a resonant structure within the antenna matched to a fraction of a wavelength of the required frequency, e.g., dipoles are a quarter of the operating wavelength in size. Whilst allowing the designer to match the antenna to the correct frequency, the quarter wave dependency imposes a bandwidth limitation on the antenna.

There are antenna designs that exploit the resonant properties and have multiple resonant structures. This gives rise to operation in a number of frequency bands. Figure 2-16 shows a graph of antenna match (or reflection coefficient) for a single-band antenna and a multi-band antenna.

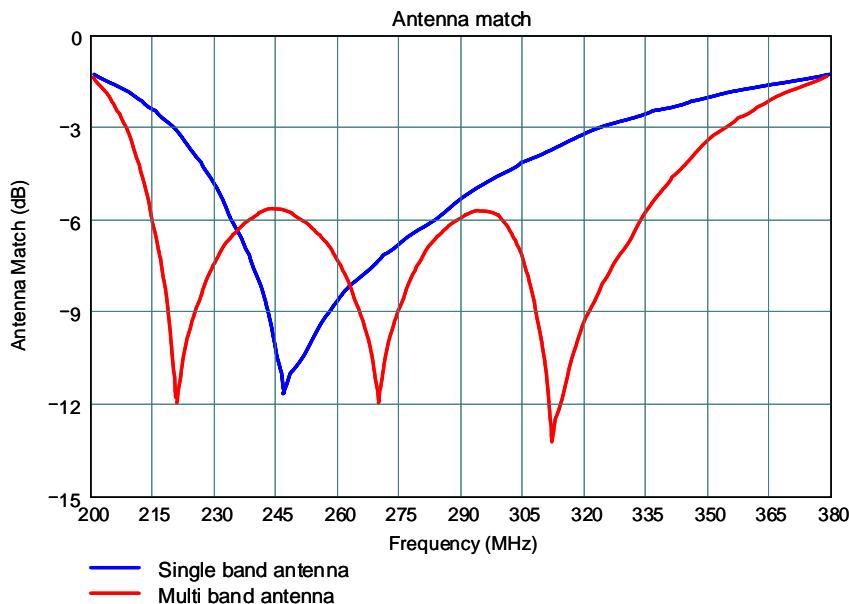


Figure 2-16: Graph showing single-band and multi-band antenna operation

The single band antenna has a match of better than -10 dB at approximately 250 MHz. The multi-band antenna has a similar match at frequencies of 220 MHz, 260 MHz and 295 MHz.

An example of an application that can require multi-band antennas are mobile telephones. For example, a global system for mobile communications (GSM) handset might be required to operate in either GSM900 (900 MHz) or GSM1800 (1.8 GHz) bands. Operation at frequencies in between is not a requirement. Another example application of multi-band antennas is that of wireless data communications devices supporting operation at frequencies defined by IEEE 802.11a (5.15 to 5.35 GHz) and IEEE 802.11b (2.4 to 2.48 GHz). Again, an antenna capable of covering the whole 2.4 to 5.35 GHz band is not required. A single, multi-band antenna would be sufficient.

UNCLASSIFIED

An example of a dual band antenna design is given in [20]. The antenna, described as being ‘F shaped’ and designed to cover the two IEEE 802.11 bands, is shown in Figure 2-17, along with a plot of the simulated return loss.

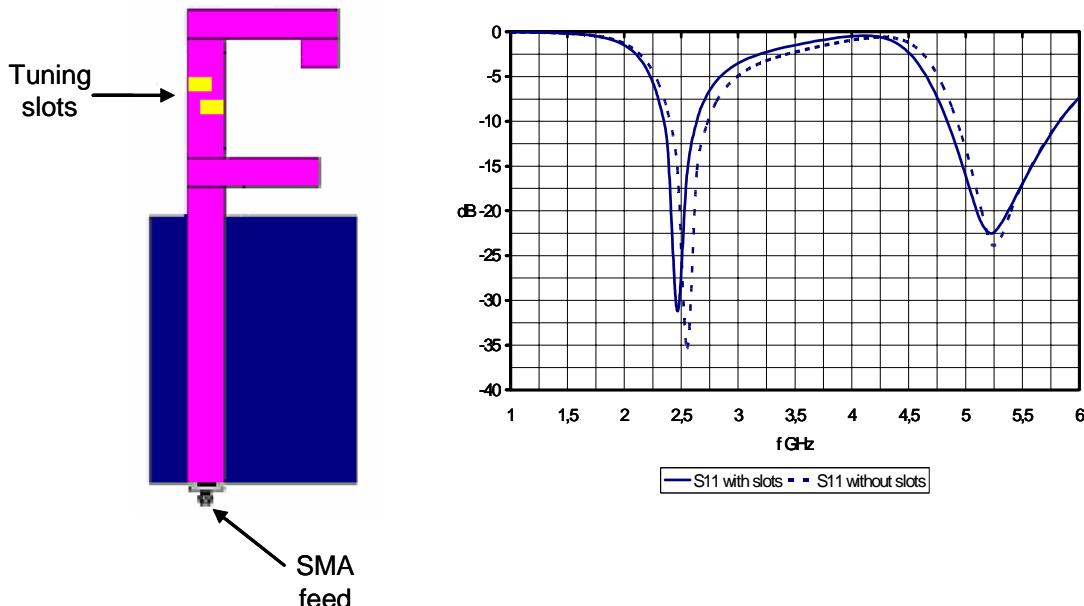


Figure 2-17: F-shaped antenna geometry and simulated return loss [20]

There are two clear resonances at 2.4 GHz and 5.4 GHz, corresponding to the IEEE 802.11 bands. The lower resonance is brought about by the length of the longest side and the top ‘hook’ of the F shape, which is designed to be approximately a quarter of a wavelength at 2.4 GHz (~40 mm). The centre ‘stub’ provides the resonance at the higher frequency. Two slots were added, the dimensions of which allow a slight tuning of the antenna. The return loss plot shows the antenna is essentially unusable between 2.7 GHz and 4.7 GHz.

The term ‘ultra-wideband’ can be misleading. Some antennas deemed to be ultra-wideband are actually multi-band antennas with a wide frequency distribution between bands of operation. These antennas are not efficient at all frequencies within this band, but are efficient at particular frequencies, e.g., the dual-band antenna shown in Figure 2-17 might be described as “a 2.4 - 5.4 GHz ultra-wideband antenna”, even though there are only two distinct frequency bands of operation.

An excellent example of a wideband antenna is the fractal element antenna (FEA) [21]. Fractals are composite geometric designs that are repeated on many scales, thus termed ‘self-similar’. In deterministic fractals, a geometric shape is repeated with a combination of scalings, rotations and translations, allowing a complex shape to be defined by only a few parameters. Figure 2-18 shows a simple equilateral triangle that has undergone four stages of transformation to form a basic fractal design.

UNCLASSIFIED

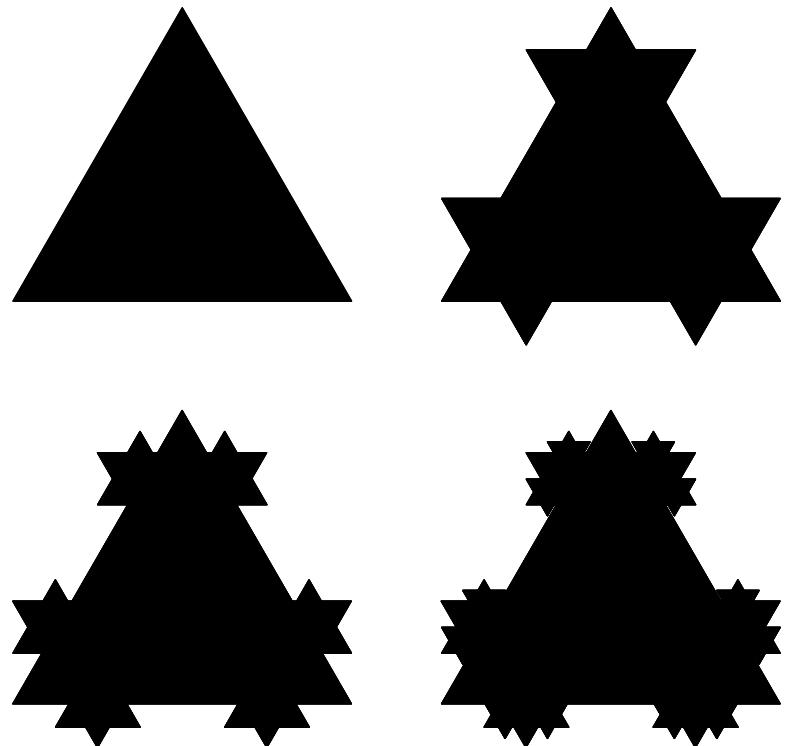


Figure 2-18: Fractal triangle antennas formed from four stages of transformation

The fractals in Figure 2-18 have the main triangle in each new stage, but with the addition of small rotated triangles in successive stages. Another arrangement, shown in Figure 2-19, removes the largest element in each successive stage, and replaces it with five smaller copies.

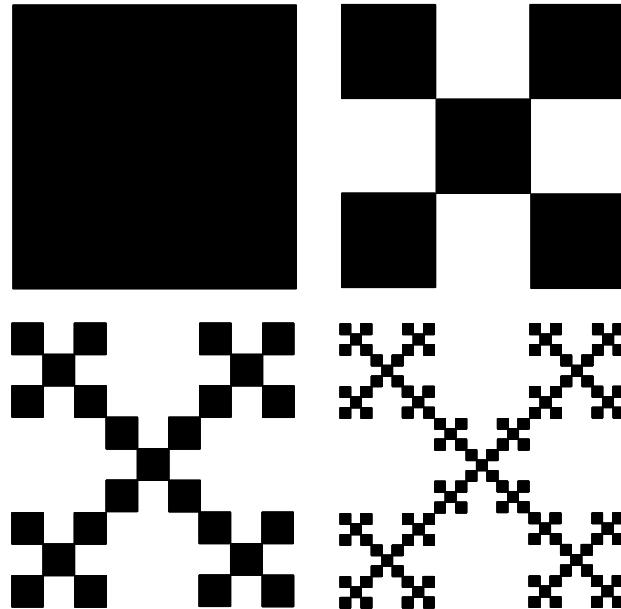


Figure 2-19: Fractal square antennas formed from four stages of transformation

One advantage of using this type of structure for an antenna is that the self-similarity yields numerous structures of resonance over a range of scales simultaneously. These can be exploited in a prescriptive way to produce the same radiating characteristics over a very broad band or, more accurately, over many distributed bands. For the antennas in Figure 2-18, the first stage triangle will have one resonance; the second stage will have two and so on. Additional advantages are:

- Compact size compared to conventional antenna designs whilst maintaining good radiation efficiencies and gains
- Mechanical simplicity and robustness. The characteristics of FEAs are obtained through their geometry and not by additional components
- FEAs may be designed for particular multi-band characteristics including specific stop bands as well as pass bands.

Fractal element antenna designs exist that are capable of producing gains of 5 dBi over passbands of 15:1, with lower gains achievable over passbands of greater than 20:1.

UNCLASSIFIED

2.4.5

Plasma Antennas

Research has been conducted, primarily by Plasma Antennas Ltd in collaboration with QinetiQ, into a new class of antenna, the solid-state plasma antenna (SSPA). This constitutes a new approach to compact, low-cost, agile beam antennas suitable for a range of RF applications. The antenna comprises a circular dielectric lens within a parallel plate waveguide structure and a small, electrically generated plasma region to form and steer beams in both transmit and receive modes.

The plasma antenna technology can currently provide:

- 360° azimuth coverage
- Very low weight and small volume
- High directivity and gain or an omni-directional configuration

Future planned developments of the antenna will enable provision of:

- Multi-octave bandwidth antenna structures (e.g., 2 to 6 GHz, 6 to 18 GHz and 18 to 40 GHz)
- Multiple polarisation capability
- Integrated low noise amplification to optimise receiver performance.

Plasma is an ionised gas that behaves as a conductor when highly energised. The plasma antenna uses this principle to generate localised concentrations of plasma within a silicon substrate to form a plasma mirror, which deflects an RF beam launched from a central feed position. This is shown in Figure 2-20. The plasma mirror is formed within a silicon wafer, metallised on upper and lower surfaces to form a parallel plate waveguide to constrain the beam. The antenna is fed centrally to form a beam through a combination of the hyperbolic plasma reflector and the circular lens properties of the wafer. A horn structure enables energy to be launched into free space.

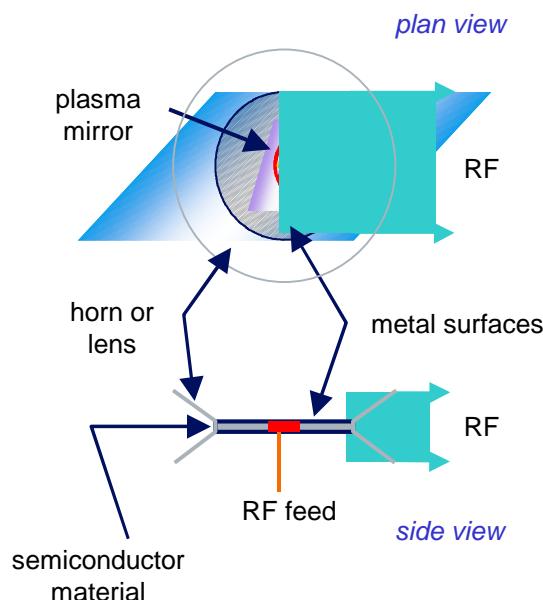


Figure 2-20: Solid state plasma antenna concept

UNCLASSIFIED

Manipulation of the position of the plasma mirror allows the beam to be steered within the plane of the semiconductor wafer. The plasma can be generated using electronically controlled devices (plasma diodes), that are positioned between the closely spaced metal surfaces.

Since the antennas are based on lens structures and reflecting surfaces, they are inherently wideband and can be used in conjunction with profiled feeds and Vivaldi-like launches into free space. This provides linear and circular polarisations across multi-octave bandwidths with good impedance matches.

The solid state plasma technology provides a steerable antenna with no moving parts, but without the need for complex beamformer networks or array processor functionality. The technology is lightweight, low volume and low cost, making it a contender for both base station and mobile handset applications. The antenna has a stackable configuration so may be operable over a very wide band and can be configured for matching beamwidths across a wide band.

2.4.6 Antenna Modelling

Developments in electromagnetic modelling software have significantly advanced antenna design in recent years. The more sophisticated software allows designers to model three-dimensional structures, such as antennas, enabling a much faster development and production of a specific design. The software has proved to be accurate when comparing real measurements from manufactured antennas to the software predictions.

Ansoft's 'high frequency structure simulator' (HFSS) software is an industry standard tool utilising a sophisticated finite element system to derive an electromagnetic field solution [22]. The problem volume is divided into smaller sub-regions (finite elements) and the field in each of these smaller regions represented with a local function. In HFSS these small sub-regions take the form of tetrahedra, with the entire collection of tetrahedra referred to as the finite element mesh.

The values of vector field quantities (such as the electric or magnetic field) at points within a tetrahedron are interpolated from the vertices of the tetrahedron. By representing field quantities in this way, the software is able to transform Maxwell's equations into matrix equations that are solved using traditional numerical methods.

Although HFSS is capable of producing a starting mesh, and subsequently refining the mesh in order to achieve a stable solution, the interpretation of a 'stable solution' is left somewhat to the user. Similarly, the level of stability which has been reached, and how this equates to the accuracy of the answer produced, is a matter for the experience of the user. In essence, HFSS is a sophisticated tool that gives highly accurate answers when used proficiently.

The rapid development of modern desktop computers has allowed designers to make full use of electromagnetic modelling software. Designers are able to model ever more complicated antenna structures using faster computers with increased memory capacity. Furthermore, antenna arrays with modest numbers of elements can now be modelled in useful timescales, and the fact that some RF components can be modelled alongside the antenna means that antennas can now be designed as part of the system.

UNCLASSIFIED

2.5 Epoch for the Development of Antenna Technology

Table 2-1 provides an indication on the epoch for the development of antenna technology; the forecast is based on the five principal antenna types identified within this chapter.

ANTENNA TYPE	ABILITY TO MEET SDR REQUIREMENTS	SUITABILITY	LIKELY COST	AVAILABILITY	IMPLICATIONS AND COMMENTS
Phased array Frequency range: 1 GHz – 10GHz	Adaptive beam steering allows spatial diversity, potential for MIMO applications	Base stations primarily, may find uses in mobile units for MIMO	Expensive due to technical complexity, some emerging low cost techniques	Expensive techniques available now (mainly military), low cost near future (civil & military)	Apart from the three main domains, (time, frequency & code) phased arrays open up spatial domain for user separation
Reconfigurable Frequency range: 5 GHz – 10GHz	Solves some issues of wideband antennas, antennas with < 10:1 bandwidth	Mobile units and base stations	Expensive initially, depending on COTS availability of MEMS technology	Likely 2008 onwards	May require sophisticated control systems to dynamically adapt
Polarisation diversity Frequency range: 200MHz – 10GHz	Ability to TX and Rx on multiple polarisations for different applications	Mobile units and base stations	Low cost printed antenna techniques employed	now	Generally requires more complex feed arrangements
Multi-band & UWB Frequency range: 500MHz –10GHz	Solves some issues of wideband antennas, antennas with < 25:1	Mobile units and base stations	Low cost printed antenna techniques employed	now	Exploitation of resonant properties gives narrow instantaneous bandwidth, but multiple resonances allow use over wider bandwidth
Plasma antennas Frequency range: 2 GHz – 40GHz	Electronically steerable, can be scaled for different frequencies	Base stations primarily	Low cost semiconductor technologies	Likely 2009 onwards	Novel immature technology, shows promise as a low cost alternative to phased arrays for some applications

Table 2-1: Epoch for Antenna Development

UNCLASSIFIED

2.6

Conclusions

The concept of SDR, by its nature, does not seem to have yielded a generic system specification. Consequently, a number of assumptions pertaining to the operating frequencies of SDR have been made. There is a significant thrust towards an SDR communications solution for the military, more specifically the JTRS concept. The bandwidth likely to be occupied by such a system is a frequency range of 2 to 2000 MHz, covering the majority of the existing bands currently in use. However, there is a growing interest in both the military and civil markets in SDR applications above 2000 MHz, even extending to about 50 GHz.

Significant headway has been made in the RF and DSP processing systems for SDR and indeed may be considered as enabling technologies. However, most of the published work on technologies for SDR applications falls short of discussing antenna options. It is not clear whether this is due to a general lack of understanding of antenna physics within the SDR community, or the recognition that producing a truly wideband antenna covering many octaves is an extremely difficult thing to achieve.

In this chapter a number of candidate antenna solutions have been summarised, including beam-steering and beam-switched phased arrays and reconfigurable and novel antenna technologies.

The aspiration of SDR to enable a single mobile handset to be configured and used for multiple roles is realisable. However, operation in *all* possible roles is not. It is more likely that SDR handsets will be designed to be configurable for groups of applications with similar antenna requirements. This will significantly ease the constraints on the antenna as specific applications may only require specific narrow bands within a much wider band to be covered. Antennas capable of this type of operation have been discussed.

Hardware technologies for implementing SDR concepts are increasing in readiness level. When considering the widest possible scope of applications for SDR, the antenna design currently provides a significant design challenge. However, state-of-the-art modelling suites and computing power allow antenna design to become an integral part of the system design. Thus, consideration can be given to antenna selection at a much earlier stage than has previously been possible.

An indication on the epochs for the development of antenna technology has been provided, based on the five principal antenna types identified within this chapter.

3 Radio Frequency (RF) Linearisation

By Markus Rupp, Technical University of Vienna.



3.1 Introduction

This chapter considers radio frequency (RF) linearisation. The power amplifier is a key element in every radio communications system, and is required to serve the function of amplifying the information-bearing signal, without distortion, in an efficient manner. However, distortion in power amplifiers occurs in two different ways [23]:

- nonlinear distortions
- linear, temporal and frequency dispersive effects.

Nonlinear distortions arise from the fact that a power efficient conversion of supplied DC power to RF signal power requires a nonlinear operation of the amplifier to a greater or lesser degree. Employing nonlinear amplification reduces the signal fidelity, so-called ‘in-band’ distortion occurs if seen in the frequency domain. Nonlinear amplification has a second major drawback. Due to the generated harmonics, out-of-band emissions occur, which can potentially cause interference to users in neighbouring frequency bands. This interference must be kept low. Note also that, due to nonlinear effects, energy is transferred into undesired bands, thus lowering energy efficiency. Pre-distortion techniques thus have the potential to make power amplifiers more power efficient, which is an important factor for handheld terminals.

The dispersive effects in power amplifiers have their origin in internal memory effects. Memory effects with long time constants (i.e., in the order of one second) occur due to thermal and trapping effects causing linear time dispersions. Memory effects with short time constants occur due to non-vanishing transit times in transistors as well as the analogue circuitry, bonding inductances, capacitances and resistances causing linear frequency dispersion of the signals.

In order to observe the dispersive effects, certain signal properties are required. If the symbol period of the signal is small compared to the inverse of the bandwidth that is to be transmitted by the power amplifier, the dispersion effects can be neglected. Nonlinear effects, however, become visible even for small signal bandwidths. The presence of both linear dispersive and nonlinear memoryless effects shows up in some unpredicted way; examples are provided in the following sections, including a discussion on these effects for modern cellular systems.

3.1.1 Software Defined Radio

Software defined radio (SDR), a movement in industry devoted to making radio transmissions all-digital and removing the analogue RF end parts entirely, started in the mid-90s. Today the SDR Forum [24] boasts more than 130 members. While the visionary ideas were the focus at the beginning, more practical aspects are today the centre of their activities. The standardisation of the software communications architecture (SCA) started in 2002.

UNCLASSIFIED

This flexible platform, developed by Communication Research Centre, Canada (CRC) and Harris Corporation, will allow protocol functionality to be changed over the air interface. It is an open architecture with Java (by CRC) and C++ (by Harris) based implementations, supported by an experimental (Posix compliant) real-time operating system (RTOS) from Wind River. The SCA can be seen as an operating environment responsible for deploying and interconnecting the signal processing objects of an SDR.

Clearly, the visionary concept of SDR can accomplish a lot more and bring advantages for terminal manufacturers, wireless service operators and terminal users. Manufacturers can concentrate research and development (R&D) on smaller hardware platforms applicable to every cellular system. This would enable development cost to be spread not over just one product but over an entire product family. Other advantages include lower production costs as a result of fewer, higher volume production lines and an ability to update and release software in a phased manner. From the point of view of the operator, new services can be implemented after initial system deployment, services can be optimised to give differentiation from other operators and multi-standard base stations would permit a wider range of services to be provided. Finally, SDR has the potential to give the end user an ability to roam across multiple cellular systems, have personal terminal configurations and increased handset lifetime before obsolescence.

The success of SDR is strongly dependent on advances of other technologies, like fast analogue/digital conversion techniques, RTOS for hardware architectures as well as low power and/or power-aware digital design techniques. Since with an SDR architecture, the hardware can be adapted to the needs of the algorithm, such devices are suitable for advanced algorithms in the digital domain, like adaptive algorithms for power amplifier pre-distortion. Thus, SDR could open a path for applying adaptive pre-distortion techniques to cellular handsets as well as base stations.

3.2 RF Linearisation and SDR

Digital pre-distortion techniques can be used to counteract the effects of both nonlinearity and linear dispersion which mostly arise from the use of analogue circuitry. These effects very much depend on the mode of operation, e.g., which modulation scheme is being used (influencing the crest factor, i.e. the peak to mean value of the signal amplitude) or the output power at which the amplifier is operating. Therefore, the benefits offered by the use of SDR, such as the flexibility of being reprogrammable and supporting various adaptive algorithms, will make it possible to operate different modes of RF transmission. This will in turn require different algorithms for RF linearisation.

UNCLASSIFIED

3.3 RF Linearisation and Mobile Communication Standards

3.3.1 Power Amplifier Linearisation in 2G Systems

2G and 2.5G systems like GSM or GSM/EDGE employ constant envelope modulation schemes (specifically Gaussian minimum shift keying (GMSK)). Further, the signal bandwidths are in the kilohertz region, which implies that the amplifier can be considered as a system without dispersion. Nonlinear amplification causes out-of-band emissions that can be reduced by means of filtering before radiation. In-band distortions caused by intermodulation are, to some extent, tolerable, since amplitude distortions will not harm the signal fidelity. If a linearisation scheme is employed in a 2G and 2.5G system, the expensive filtering of the amplified signal can be omitted and the nonlinear distortions of the phase (and amplitude) can be corrected. Since the power amplifier appears as a dispersion-free system, the broadband linearisation is not required. In 2G (2.5G) systems analogue linearisation schemes (e.g., feedforward linearisation [25], [26]) are used. Digital pre-distortion is also used, applying simple schemes like look-up table methods [27], which compensate for a static (memoryless) nonlinearity. Note that pre-distortion techniques in 2G have only been used in base stations, where complexity and power consumption is less of an issue. For cellular handsets, low cost demands the use of passive filtering rather than additional ADC conversion, down mixing and DSP algorithms.

3.3.2 Linearisation in 3G and Beyond

Third generation systems like UMTS (Universal Mobile Telecommunications System) and Wideband-Code Division Multiple Access (W-CDMA) and currently investigated systems like ultra-wideband (UWB), employ spectrally-efficient linear modulation schemes such as M-ary quadrature amplitude modulation (QAM), where not only the phase of the signal bears information, but the amplitude of the carrier is modulated by the information bearing signal. The signal bandwidths have increased drastically also. In UMTS, one channel occupies 5 MHz. These facts demand linear amplification over a wide range of frequencies. Although simple linearisation schemes, like feedforward linearisation and digital pre-distortion based on a static model of the power amplifier can remedy nonlinear effects to some extent, they fail to compensate for linear dispersion. Digital pre-distortion based on a dynamic power amplifier characterisation is a promising candidate for future linearisation schemes. Note that, however, such algorithms are of much higher complexity than the static techniques of 2G systems and are thus typically considered for base stations only.

It is very important to understand that, when considering power amplifier behaviour, such linear and nonlinear effects cannot be simply separated. For communication with signals of relatively small bandwidths in the kilohertz region (as in second generation (2G) and 2.5G systems) only nonlinear memoryless effects occur.

However, third generation (3G) systems employ signal bandwidths in the megahertz region, e.g., 5 MHz in a single carrier wideband-CDMA (W CDMA) system (3.84 Mchips/s, root-raised cosine filtering with a roll-off of 22%). These signals are distorted not only by nonlinear effects but also by the linear dispersive effects of the power amplifier. Like in a dispersive communication channel, (linear) equalisation at the transmitter, the source of the dispersion, can remedy these.

UNCLASSIFIED

Modern devices with even higher bandwidth (for example, WLAN (Wireless Local Area Networks) 802.11n may require a 40 MHz bandwidth), as well as the necessity to combine several channels for amplification, complicate the problem of nonlinear distortion including memory effects even more.

The required linear amplification of the information bearing signals through power amplifiers is accomplished by using a linearisation scheme either working in the analogue, RF part or the digital, baseband or intermediate frequency (IF) part of the transmitter. Analogue linearisation schemes are either not able to compensate for the linear dispersion effects or are only able to do so to a limited extent. Digital pre-distortion techniques, however, can be used to counteract the effects of both nonlinearity and linear dispersion.

3.4 Power Amplifier Linearisation Methods

In this section we will review the most common linearisation methods for power amplifiers (PAs). These methods are split into analogue methods that work entirely in the RF domain, or digital methods, which try to compensate for nonlinear and dispersion effects in the digital domain using DSP hardware.

We will compare the presented methods with respect to their ability to compensate nonlinear effects, as well as their performance in compensating for dispersion effects. The key hardware requirements of each method are also investigated. The hardware requirements determine largely the integration feasibility in an SDR system and its implementation costs.

3.4.1 Analogue Methods

Analogue methods work exclusively in the RF domain. Since SDR technology tends to reduce the costly analogue part in the transmitter/receiver chain, these methods will not be candidates for such systems. Nevertheless, the performance of digital pre-distortion has to be compared with the more traditional analogue methods and thus a short overview on such methods is provided.

3.4.1.1 Feedforward Linearisation

A simplified model of a feedforward linearisation scheme is shown in Figure 3-1. A small portion of the signal at the output of the nonlinear PA is compared with the RF signal at the input to the PA. The resulting ‘error’ signal contains only intermodulation distortion products caused by the nonlinearities of the PA. These signals are amplified and inverted by an error amplifier. If the distortion products are in opposite phase, the summation of the delayed main signal in the upper branch and the distortion products in the lower branch yields a signal with less distortion being delivered to the antenna. Input and output of this scheme are in RF domain.

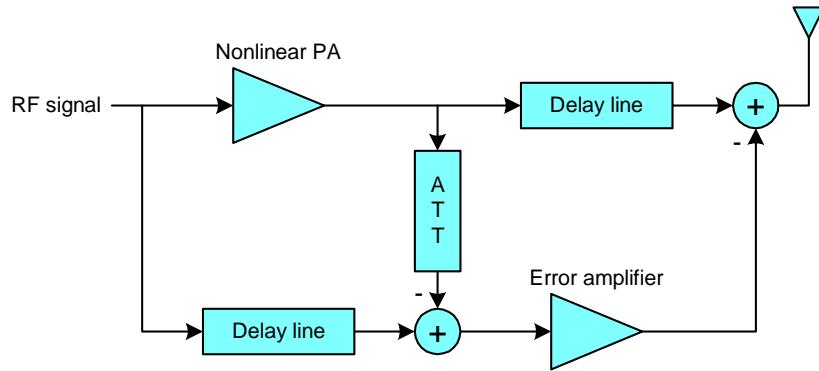


Figure 3-1: Feedforward linearisation

The feedforward scheme is an open-loop configuration and is, therefore, vulnerable to drift effects. On the other hand, it is unconditionally stable (as there is no feedback path). The main drawback of this method is the high demands on gain and phase matching of the various components. It is difficult to maintain this matching over temperature variations, time, amplifier loading, or to the change in power levels if the number of carriers change in a multicarrier application. Furthermore, the delay lines in the main signal path and the error amplifier reduce the efficiency of the technique [25].

3.4.1.2 Feedback Linearisation

A block diagram of a Cartesian loop linearisation scheme is shown in Figure 3-2. A portion of the output signal is fed back to the input. It is down-converted and demodulated before these quadrature components are subtracted from the input signal to build a loop error signal, which, after filtering, drives the modulator and the nonlinear RF power amplifier. The bandwidth of the loop must be sufficiently wide to accommodate all significant intermodulation products of the power amplifier. Since the scheme is a closed-loop configuration it is potentially unstable. The stability issue is of a very difficult nature since the loop includes both linear and nonlinear components and very few existing theoretical results for such systems are applicable. On the other hand, due to the controlling nature of such a scheme, it can compensate for drifts in power amplifier nonlinearities caused by temperature changes, DC power variations, load changes, and component aging.

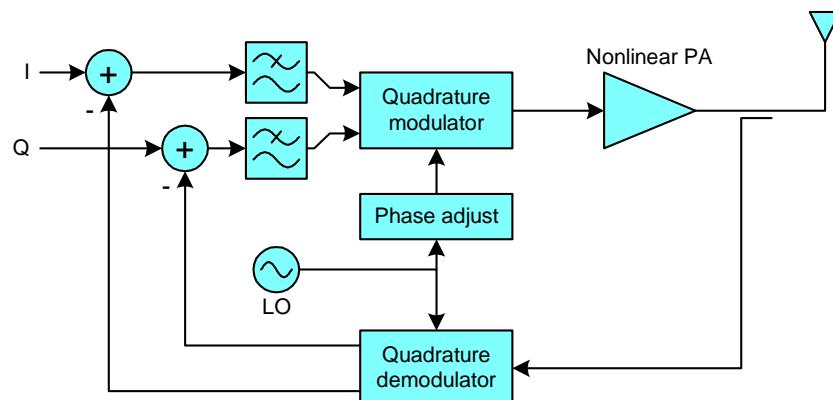


Figure 3-2: Cartesian feedback transmitter

UNCLASSIFIED

The feedback configuration is able to reduce the intermodulation products of the nonlinear power amplifier by a factor equal to the gain of the loop. However, a large gain implies a low stability margin. Furthermore, the noise performance is no longer determined solely by that of the RF power amplifier but by the loop as a whole [28].

3.4.2 Digital Methods

Figure 3-3 shows a simplified block diagram of a digital pre-distortion scheme with crest factor reduction. Both the crest factor reduction and digital pre-distortion operate in the digital domain. The crest factor reduction aims to reduce the high ‘crest factor’ of modern communication systems like multi-code (MC) MC-CDMA and orthogonal frequency-division multiplexing (OFDM) in order to ease the subsequent pre-distortion task. After the pre-distortion block, a DAC converts the information bearing signal into an analogue signal, which is then up-converted to RF.

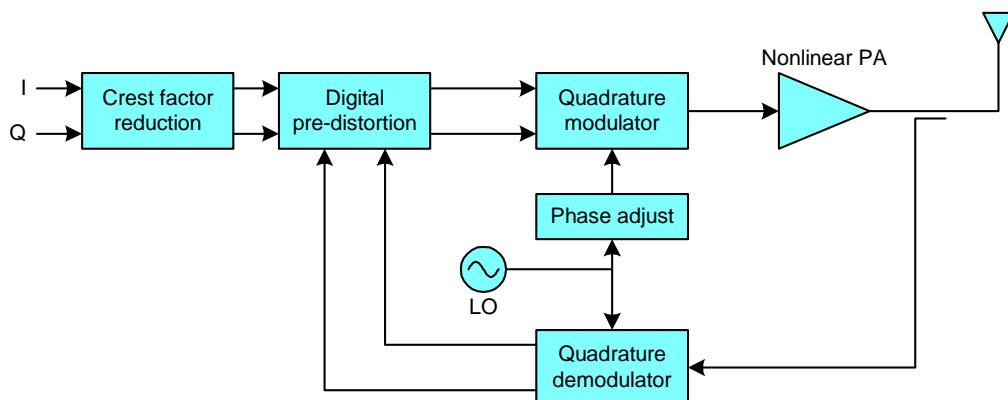


Figure 3-3: Digital pre-distortion with crest factor reduction

The digital pre-distortion method is a feedback scheme and therefore able to react to drifts of the nonlinear power amplifier. Two different approaches are commonly used for determining the pre-distortion filter:

- Inverse Modelling
 - Direct Inversion.

We will now consider each of these in turn.

3.4.2.1 Inverse Modelling

The inverse modelling scheme is depicted in Figure 3-4. Here, by observing the output of the PA, the algorithm aims to find the post-inverse, which, when concatenated with the PA, would act as a linear system. Since the digital algorithm can only be placed in front of the power amplifier, the post-inverse is copied as a pre-inverse [29], [30]. The inherent problem with this scheme is that the post-inverse and pre-inverse are not identical (only in the case of linear, time-invariant systems) and the performance is thus seriously limited.

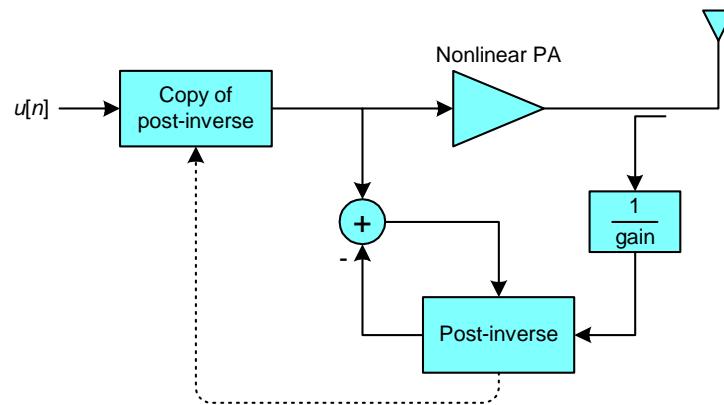


Figure 3-4: Inverse modelling

On the other hand, the pre-distortion unit is identified directly, without the need for a model of the power amplifier, which further reduces the complexity.

3.4.2.2 Direct Inversion

The direct approach, shown in Figure 3-5, is more complex than the inverse modelling approach in that first a model of the PA has to be identified. Based on that model, the pre-distortion filter has to be constructed. The problem is thus twofold – modelling and inversion of a nonlinear dynamic system.

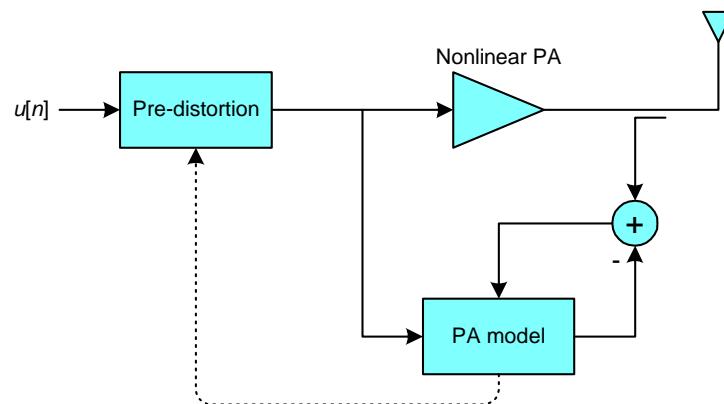


Figure 3-5: Direct inverse

3.4.2.3 Crest Factor Reduction Techniques

In modern communication systems, the transmitted signal is a superposition of multiple orthogonal signals (e.g., Walsh-Hadamard coded signals in CDMA or narrowband signals at distinct frequencies as in OFDM). This, together with a non-constant envelope modulation format (e.g., QAM), causes a high crest factor of the transmit signal, which deteriorates the signal fidelity in the presence of (weakly) nonlinear devices, such as power amplifiers. The crest factor (CF) is defined as:

$$CF = \sqrt{PAPR} \quad \text{Equation 3-1}$$

where $PAPR$ is the peak-to-average power ratio (PAPR), i.e.:

$$PAPR = \frac{\max_{0 \leq t \leq T} |s(t)|^2}{\frac{1}{T} \cdot \int_{t=0}^T s^2(t) \cdot dt} \quad \text{Equation 3-2}$$

We note that the PAPR is simply the square of the crest factor. Typical values for crest factors are reported in [31] and listed here in Table 3-1. Note that a simple sinusoid already has a crest factor of 3 dB. It is thus not likely to obtain smaller values. An M -ary phase-shift keying (PSK) is expected to preserve this 3 dB. However, due to filtering at the transmitter, the true values are often larger (e.g., 3.21 dB for 8-PSK).

Technology	CF
W-CDMA (128 occupied channels, RRC filtering with 22% roll-off)	13.6 dB
16 QAM (symbol rate 25ksymbols/s, RRC filter with 20% roll-off)	6.03 dB
GSM EDGE (8-PSK)	3.21 dB

Table 3-1: Crest factors for common modulation formats

Since OFDM and CDMA are mostly favoured in modern communication systems, the research activity is mostly focussed on these, leading to very different methods.

The methods in OFDM can be categorised into two main approaches:

- Coding methods
- Clipping methods.

UNCLASSIFIED

Both methods are targeted to work in the digital domain, i.e., before DAC conversion. The principle of the coding methods is to reduce the crest factor through the careful selection of the code words [32], e.g., by excluding code words with a high crest factor [33]. This has the negative side effect of reducing the code rate. Other methods modify the code words by adding a set of phase shifts to form modified code words with a lower crest factor [34][35]. The error correction capabilities, code rate, encoding and decoding complexity are preserved. With these methods the crest factor can be reduced typically by between 3 and 4 dB.

Clipping techniques attempt to reduce the crest factor by limiting (i.e., ‘clipping’) high signal peaks [36] [37]. Nonlinear distortion in power amplifiers can be reduced in this way, but the clipping of signal peaks introduces nonlinear distortion itself. This can be alleviated by subsequent filtering. Clipping and filtering can be repeated, since the smoothing after one clipping stage can increase the signal peaks once more [38]. Typically, with clipping techniques, the crest factor can again be reduced by between 3 and 4 dB.

The problem of a high crest factor in MC-CDMA systems has been studied recently in [39][40][41][42][43][44][45][46]. Here, Reed-Muller codes, Walsh-Hadamard transform (WHT) and Bent Functions for MC-CDMA [39] [40] have been applied. Special interest has resulted in the carrier interferometry (CI) codes [41][42][43][44] as well as complementary-sequence-based MC CDMA signals [45][46].

3.4.2.4 Static Pre-distortion vs Dynamic Pre-distortion

Whether a memoryless or a model with memory is more suitable for a given power amplifier strongly depends on the signal bandwidth. Broadband signals, e.g., MC-CDMA signals, require a power amplifier model with memory to additionally accommodate for the dispersion introduced [47]. The pre-distortion filter is thus required to contain memory as well.

Figure 3-6 shows a measured amplitude modulation (AM)-AM conversion (amplitude of the output signal $|y(n)|$ vs amplitude of the input signal $|u(n)|$) of a standard power amplifier (Minicircuits ZLH-42W) [48]. The input signal is a multi-tone signal with a bandwidth of 2 MHz. Memory effects cause the characteristic to broaden; noise is negligible due to the usage of high performance measurement equipment. For purposes of this test, the amplifier was driven hard into saturation.

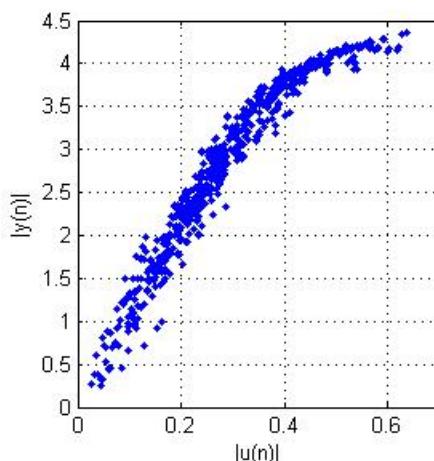


Figure 3-6: Measured AM-AM conversion

UNCLASSIFIED

A simulation result (obtained using Agilent's advanced design system (ADS)), utilising a W-CDMA signal, is depicted in Figure 3-7, where again the AM-AM conversion is shown. The power amplifier is driven in a weak nonlinear fashion. Again, the characteristic is broadened, which indicates a system with dispersion.

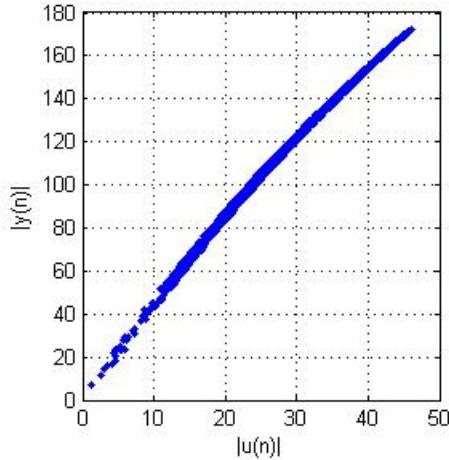


Figure 3-7: Simulated AM-AM conversion with a W-CDMA signal

Static models and the resulting static pre-distortion filters (e.g., based on a look-up tables [27]) yield poor results. Models incorporating memory effects are mandatory. Common models for nonlinear systems are based on Volterra series [49][50][51] and simpler structures, such as Wiener and Hammerstein models [48][52][53][54].

In Figure 3-8, the pre-distortion is based on the direct inverse approach [55] and assumes a memoryless characteristic of the power amplifier. Based on this static characteristic, the pre-distortion unit can only correct for the nonlinear behaviour, i.e., the dispersion cannot be compensated. Since the amplifier is operated strongly in saturation, a lower gain has to be accepted. Note that in practical applications, the power amplifier is driven in a more linear fashion compared to Figure 3-7, which eases the pre-distortion task and no loss in gain occurs.

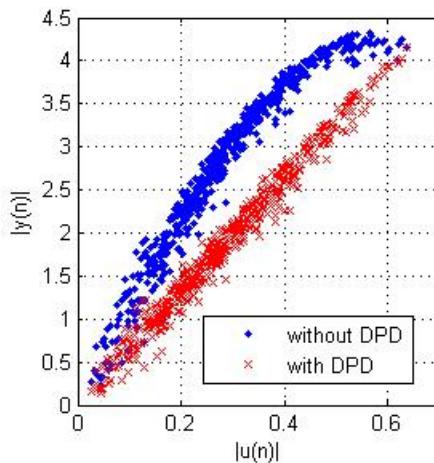


Figure 3-8: Digital pre-distortion with static power amplifier model

UNCLASSIFIED

Using a dynamic power amplifier model (which leads also to a dynamic pre-distortion filter), dispersion as well as the nonlinear effects can be compensated. Figure 3-9 shows the results. The model used for the power amplifier is a 5th-order Volterra series. When compared to Figure 3-8, both pre-distortion methods (static as well as dynamic) cause improvements. However, as the much finer line in Figure 3-9 indicates, the linearity achieved with the dynamic pre-distortion method is significantly better. The performance improvement can be measured in terms of error energy.

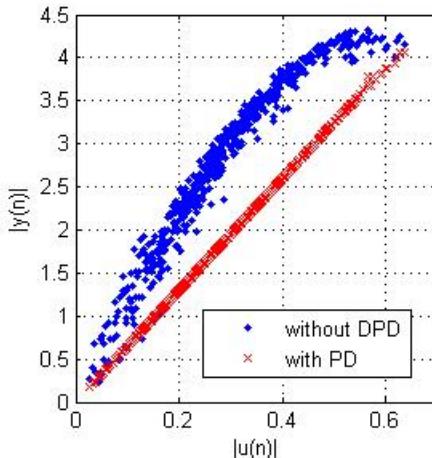


Figure 3-9: Digital pre-distortion with dynamic power amplifier model

The relative errors achieved with both methods are reported in Table 3-2, where the relative error is computed as:

$$\varepsilon = 20 \cdot \log \left(\frac{\|d(n) - y(n)\|_2}{\|d(n)\|_2} \right) \text{ dB} \quad \text{Equation 3-3}$$

where $d(n) = g \cdot u(n)$ is the targeted signal (with a gain of $g = 6.5$), $u(n)$ is the input signal, and $y(n)$ is the signal with or without pre-distortion.

	Static Pre-distortion	Dynamic Pre-distortion
Relative error, ε	-20.5 dB	-44.8 dB

Table 3-2: Relative errors for static vs dynamic pre-distortion

Correcting for the dispersion yields large performance gains in terms of relative error. In our example, more than 20 dB are gained, which clearly shows that simple pre-distortion algorithms based on a static power amplifier model are, in wideband communication systems (particularly if more carriers are to be amplified as in MC-CDMA systems), unsuitable.

3.4.2.5 Non-Adaptive vs Adaptive Pre-distortion

Both methods presented, inverse modelling and direct inversion of a power amplifier model, need to be worked in an adaptive form. In a technical implementation (e.g., using a DSP or FPGA) least mean square (LMS)-type algorithms [56][57] are commonly used for the parameter identification and adaptive tracking of system variations. Recursive least squares (RLS) algorithms are poorly suited for implementations using fixed-point arithmetic (at least in simple multiplier/accumulator (MAC) forms of DSPs). LMS-type algorithms are more robust [58][59] and less complex for implementation. The complexity in terms of multiplications grows linearly with the number of parameters to estimate whereas for RLS, the complexity of a somewhat robust form of the algorithm is typically of quadratic order. Note that solutions with linear complexity exist but are not likely to be numerically stable in fixed-point implementations.

Adaptive identification of Volterra models (i.e., direct approach) or post-filters (i.e., inverse modelling approach) is identical to the traditional adaptive identification of linear filters. Wiener- and Hammerstein models require the identification of two filters, one a linear filter, the other a static nonlinearity. Analysis of the adaptive gradient algorithms follows the same lines as for linear adaptive filters [60][61]. Next to the classic LMS and RLS algorithms, many variants exist (in literature more than 100 are reported), offering improvements for specific problems. All algorithms have in common that they are minimising a cost function, or at least approximate such a minimisation.

The error surface, i.e., the error energy as a function of the desired parameters, is typically multimodal in the case of nonlinear power amplifier problems. Thus, gradient based algorithms tend to find only local minima and may entirely neglect the global optimum. While such multimodal cost function problems are typically solved by genetic algorithms or simulated annealing, algorithms that at least have the potential to find global optima exist. Note that, however, even such procedures cannot guarantee to find the global optimum. They often require a high complexity that cannot be predicted beforehand.

Thus, newer methods not based on gradient methods may be of more interest. Recently, [55] has formulated the problem as a zero-finding procedure and has obtained excellent results.

In Figure 3-10, the equalisation quality of the method proposed in [62] is illustrated. The nonlinear system to equalise is a fifth-order Volterra system. The parameters of this Volterra system are estimated using least-squares estimation. The input and output signals are obtained via a measurement of a standard power amplifier (Minicircuits-ZLH 42W) [55]. The input signal is a multi-tone signal with a bandwidth of 2 MHz.

UNCLASSIFIED

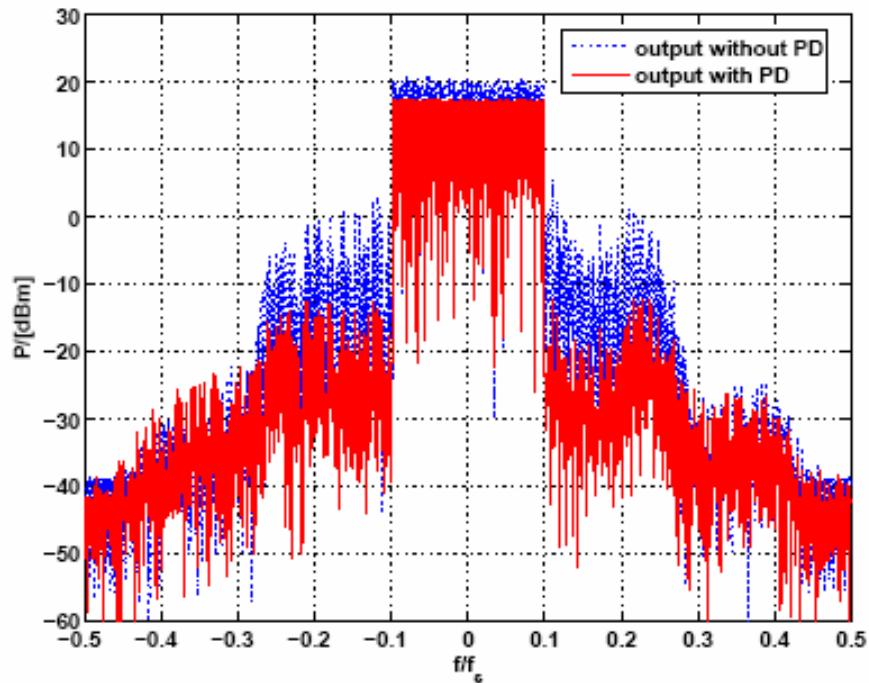


Figure 3-10: Equalisation with successive approximation [55]

Four iterations were performed with each method. Using the Newton method [63] (shown in Figure 3-11), nearly perfect equalisation is possible, whereas the successive approximation method cannot compensate for the nonlinear distortions introduced. The third-order products are significantly higher (up to 20 dB greater) and the fifth-order products are not compensated for at all using the successive approximation method. To compensate for these distortions, more iterations would be necessary, resulting in a higher complexity. It is noted, however, that the cost per iteration is lower than the cost per iteration with the Newton method.

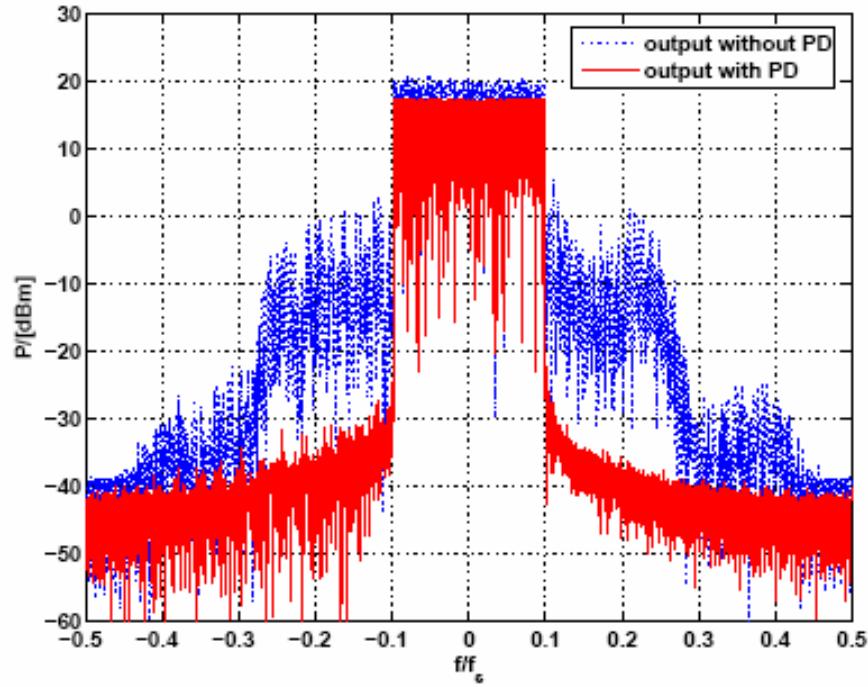


Figure 3-11: Equalisation with Newton method [55]

3.5 Technology

3.5.1 ADC/DAC Performance Overview

Data converters are a key element in an SDR system. The fundamental requirements are high resolution (i.e., the number of bits), a high sampling rate, high linearity and, for portable devices, low power consumption [64][65]. Fast and accurate ADCs up to 210 MHz with a resolution of 12 bits [66] are available. The linearity requirements for a digital pre-distortion system are a spurious-free dynamic range (SFDR) of 80 dBc, a differential nonlinearity (DNL) of ± 0.3 least-significant bit (LSB) and an integral nonlinearity (INL) of ± 0.5 LSB (typical values).

Table 3-3 gives a short overview on published results on very high-speed, high-resolution ADCs. The resolution of these high-speed converters is relatively low; eight bit resolution yields at best a SNR of 48 dB.

f_s	Resolution	Power	Technology	Reference
2 GHz	8 bits	3.5 W	SiGe	[67]
1.3 GHz	6 bits	600 mW	CMOS	[68]
1.6 GHz	6 bits	340 mW	CMOS	[69]
1.1 GHz	6 bits	300 mW	CMOS	[70]

Table 3-3: Overview of example very high speed ADCs

UNCLASSIFIED

The digital-to-analogue (DAC) converters are the other parts of the interface with the analogue front-end. DACs with a resolution of 16 bits and a sampling frequency of 600 MHz are available, e.g., the Analog Devices AD9726 [71]. This device, with an intermodulation distortion (IMD) of greater than 80 dBc, fulfils the high linearity and resolution requirements, together with a very high sampling rate.

Table 3-4 lists a few of the currently available high-speed DACs.

f_s	Resolution	Linearity	Technology	Reference
1 GSps	10 bits	SFDR > 60 dBc	CMOS	[72]
1 GSps	6 bits	-	CMOS	[73]
600 MSps	16 bits	SFDR = 80 dBc	CMOS	[71]

Table 3-4: Overview of example very high-speed DACs

3.5.2 Digital Signal Processing Hardware in SDRs

While commercial cellular radio systems for handhelds (e.g., GSM, IS-136, DECT) have used general purpose DSPs in the past, base stations have typically used DSPs and FPGAs (e.g., from Altera and Xilinx) to support flexibility. In the past, handheld terminals have remained unchanged during their life time, while base stations are updated often, even on a weekly basis. Such changes were due to updates in standards and bug fixes. In WLAN, devices are much cheaper, making changes after product delivery impossible.

However, in the future, SDR techniques will offer new flexibilities even in the hardware architecture. Surprisingly, most activities are in start up companies, with universities remaining more in the background. The following activities can be distinguished.

3.5.2.1 Fully Fixed Architecture

As already mentioned in the introduction, in the near future, SDR will offer flexibility in the domain of higher layer functions (i.e., communication protocols) using the SCA standard. The SCA standard requires a so called ‘fully fixed’ architecture, i.e., the architecture is fixed during run time and flexibility can only be achieved by software.

The flexibility of such fully fixed architectures lies only in software running on standardised processors and embedded DSPs. The hardware part of the architecture is made entirely from standard blocks, which can neither be changed at design-time nor at run-time. In case no suitable intellectual property (IP) modules are available, custom building blocks have to be first generated.

UNCLASSIFIED

3.5.2.1.1 *Sandbridge Technologies*

This USA-based company [74] claims to provide a complete SDR solution with their SB9600 processor. This processor is able to support GSM, GPRS and W-CDMA systems and is based on dual-core architecture, comprising a proprietary DSP core called SandBlaster and a standard ARM microcontroller core. In addition to these blocks, the processor also contains, among others, IP blocks for Bluetooth wireless technology, WLAN, IrDA, UART, USB, GPS, SIM, MP3, MPEG4 and JPEG functions. Since the hardware part of this architecture is fully fixed, the processor is delivered as already manufactured silicon, in 0.13 µm CMOS technology running with a 600 MHz clock speed. The only remaining flexibility lies in the software part of the architecture, which is user developed and can be modified at run-time. The design process for the software part of this architecture is derived from a system definition in MATLAB, which requires manual porting to fixed-point C code, which is in turn translated automatically by Sandbridge software tools to DSP and microcontroller code for the SB9600 processor.

3.5.2.1.2 *Intrinsity*

Another USA-based provider of SDR solutions is Intrinsity Inc [75]. The FastMATH [76] adaptive signal processor is claimed to be a suitable platform for SDR implementations. This processor comprises a million instructions per second (MIPS) central processing unit (CPU), a matrix math coprocessor, multi-level memory structures and so called RapidIO-ports. These ports are used for off-chip communications at up to 4 Gb/sec.

In most respects, this dual-core architecture is very similar to that of the SB9600 processor, with the relative advantages of a higher clock rate (up to 2 GHz, for both cores) and efficient off-chip communications and multi-chip scaling through RapidIO-ports. However, the FastMATH adaptive signal processor still suffers from the same basic weakness with respect to SDR solutions as the SB9600. It lacks run-time hardware adaptability and has a power consumption far beyond what's reasonable for a mobile device.

3.5.2.2 Flexibility at Design Time

The flexibility of the following architectures comes from both hardware and software at design time and from software at run-time. The hardware part of the architecture is composed of standard blocks, as well as custom blocks defined at design time.

3.5.2.2.1 *Application Specific Instruction Set Processor*

Brakensiek et al [77] from the Nokia Research Centre in collaboration with the University of Dortmund, Germany proposed an approach in which similarities and differences between waveform standards are identified and parameterised. To overcome some of the practical implementation issues, they moved away from 'ideal SDR' where a single hardware device is reconfigured to any standard. Instead, their software reconfigurable radio (SRR) uses as little standard-specific hardware as possible and maximises the use of parameterised modules that can implement two or more waveforms.

UNCLASSIFIED

The algorithm specific instruction set processor (ASIP) is an independent reconfigurable hardware accelerator that can be used by the DSP or microprocessor. The goal is to allow the DSP to perform complex high-speed processing on the ASIP. The ASIP itself may consist of one or more processing elements (PEs). Each PE can be designed for a particular class of algorithms, such as Galois field arithmetic, linear transformations or orthogonal transformations. The system can contain one or more PEs of a similar kind based on the needs of the system. The reconfigurable sections are specialised for certain algorithms, which allows the reconfigurable sections to be optimised for wireless applications.

3.5.2.2.2 FPFA

Researchers at the University of Twente in the Netherlands have proposed a heterogeneous reconfiguration architecture along with a QoS-driven operating system for future low-power mobile multimedia systems [78]. The heterogeneous system-on-chip (SoC) consists of a general purpose processor (an ARM core), a bit-level configurable processor (an FPGA) and several word-level reconfigurable array tiles, so called field-programmable functional arrays (FPFAs). Each of these modules is targeted for specific applications with programmable runtime configuration capabilities. For instance, the ARM is targeted for control applications like if-then-else statements and for/while loops, the FPGA for bit-level operations like pseudo number (PN) code generation and the FPFA for inner loop computations and computationally complex DSP algorithms that have a regular structure.

3.5.2.2.3 Adelante Technologies

This Belgium-based provider of embedded DSP solutions [79], a former spin off from IMEC, offers its Galactic DSP technology as a suitable platform for the implementation of SDR systems. The Galactic DSP system comprises three elements, namely, the Adelante DSP core, application specific accelerators and the ‘Atmosphere Development Environment’. Whilst the DSP core provides the standard CPU functionality on this platform and subsequent flexibility through software, of real interest to SDR implementations is the hardware flexibility found in the application-specific accelerators. These blocks can be specified at design-time to provide any required functionality in hardware, which provides a higher level of hardware flexibility than just pre-defined IP blocks.

There are two types of application-specific accelerators, the application-specific execution units (AXU) and the application-specific coprocessors (ASCP). AXUs are more tightly bound to the DSP core (they are, in fact, connected directly to the core’s internal bus). However, they are much smaller (less than 1K gates) than ASCPs (several 10K gates), which are more independent and do not require intensive communications with the DSP core. The Atmosphere Development Environment is a set of tools that help the designer develop the software for the Galactic platform. All tools are based on C and include a compiler, debugger, profiler, simulator and other support.

UNCLASSIFIED

3.5.2.3 Flexibility at Run Time

The flexibility of these architectures, described in the preceding section, comes from both software and hardware, both at design-time and, in particular, at run-time. Efforts for the flexibility at run time are described in the following.

3.5.2.3.1 *Parameterised Approach*

One of the very few University-driven approaches is the so called ‘Parameterised Approach’ by F. Jondral (see, for example, Chapter 8 in [80]). In this approach, all information from standards must be known at design time. Then, a flexible structure is generated to run all of the standards to be supported.

This approach has two major drawbacks. The first drawback is that the structure cannot be changed afterwards, so new standards cannot typically be supported. The second drawback is that, due to the parameterization, there are always some parts of the building blocks that are unused since such parts are not required by selected parameters. For example, the UMTS standard (Release 99) currently supports only quadrature phase-shift keying (QPSK) modulation. Since the modulation block is modular and can support other modulation schemes like 8-PSK and 64QAM, large parts are active but not in use. While a power down mode is easy to provide when entire blocks are switched off, it is difficult to provide when only parts of the building blocks are in use. Thus, the hardware structure will always utilise all of the available hardware regardless of the selected standard, and will thus drain maximum energy from the batteries. Therefore, the overhead in terms of chip area can be significant, resulting in uneconomic, large designs with high manufacturing costs and low yield. For cellular systems in particular, such an approach is not very desirable.

3.5.2.3.2 *Maia Reconfigurable Processor*

The Maia [81] reconfigurable processor is the outcome of a research project at the University of California, Berkeley. This processor incorporates an embedded ARM processor, arithmetic operations blocks, arithmetic logic units (ALUs) and MACs, as well as a FPGA block and an interconnect network binding all the above blocks.

Because of its research background, the Maia processor is not a commercially available platform for development of SDR implementations. However, this processor practically demonstrates hardware flexibility at run-time to reflect changing system functionality and is thus, in principle, suitable for implementing SDR systems.

3.5.2.3.3 *GARP*

Another University’s research outcome is GARP from the University of California, Berkeley [82]. GARP was designed as a configurable accelerator for general purpose processors. The designers recognised that reconfiguration times and low data bandwidths have deterred designers from using reconfigurable computing, especially in complex high-speed processing applications. GARP combines a single-issue MIPS processor with reconfigurable hardware.

UNCLASSIFIED

The system, designed for a clock speed of 100 MHz, uses rapid reconfiguration with just a few cycles of overhead and direct access to memory from the reconfigurable core. GARP's reconfigurable hardware attaches to the main MIPS processor as a co-processor. Explicit processor move instructions transfer data between the two parts. GARP is a two-dimensional array of configurable logic blocks (CLBs) connected by programmable interconnects, similar to an FPGA. However, the global clock speed remains constant. The array's four 32-bit data busses and 32-bit address bus can be used to load configuration bits and transfer data while the array is idle, and can be used to access memory while the array is active.

3.5.2.3.4 *PicoArray*

One of the very recent additions to the embedded platform market is the picoArray from the UK-based company picoChip [83]. The picoArray is in fact a heterogeneous array of smaller 16-bit processors. There are four types of array element processors. The first type is aimed at spread/despread functions as well as forward error-correction. The second type is designed for filtering, with a dedicated MAC unit. The third type is aimed at control intensive functions, while the fourth is used for scheduling operations.

The array typically consists of several hundred elements, with each processor operating using a 160 MHz clock. A rough performance guide provided by picoChip states a 20-fold performance gain of the picoArray, running at 160 MHz, over a top-of-the-line Texas Instruments DSP, running at 600 MHz. Aside from the claimed performance gains and optimisation of the array elements for wireless algorithms, the picoArray architecture has a particular suitability to future SDR implementations, offering hardware flexibility at run-time. Although each of the array elements is not itself reconfigurable, the mapping of the algorithm over the hundreds of processors in the array can change tens of thousands of times per second, thus providing flexibility in the way the hardware appears to the software components.

3.5.2.3.5 *RAW*

Massachusetts Institute of Technology's (MIT's) RAW processor [84] comprises of 16 identical programmable tiles, each consisting of one static communication router, two dynamic communication routers, an eight-stage in-order single MIPS style processor, a four stage floating point unit, a 32 kByte data cache, and 96 kBytes of software-managed instruction cache. Each tile connects only to its four neighbours. The RAW architecture exposes the interconnection between computational units inside a microprocessor to the instruction set. The instruction set architecture is designed to indicate to the programmer where the instruction would be physically executed and the number of hops between tiles necessary for the operation. This allows the programmer (or the compiler) to control the data transfer between tiles as is done in an application-specific integrated circuit (ASIC). The RAW operating system used with the tiled architecture allows for spatial and temporal multiplexing of tasks. This dynamic software control of computational resources allows for run-time reconfiguration with ASIC-like latencies.

UNCLASSIFIED

3.5.2.3.6 Quicksilver Technologies

Very similar to the picoArray is another array of processors called the adaptive computing machine (ACM) made by Quicksilver Technologies [85], based in the US. This array is specifically designed for run-time adaptability. It is built in a lattice structure, with four types of processor elements. The first type is responsible for arithmetic operations, the second for bit-manipulation operations, the third for finite state machine implementations and the fourth, a so-called scalar node, is used for the execution of legacy code. There are several important differences between the approaches of the picoArray and the ACM. The ACM is much coarser-grained than the picoArray, offering only a 32-node array and announcing a 64-node solution. However, each of the nodes in the ACM can handle more complex algorithms independently.

Also, each of the nodes in the ACM is individually reconfigurable, with the entire processor being reconfigured up to 40,000 times per second. The proprietary extension to the C language, called SilverC, is the basis of the design flow for the ACM. This design flow completely eliminates the need for hardware design, because the reconfiguration of each processor node happens at run-time, to suit the compiled software. Moreover, the design flow can even start at a higher design level utilising a Simulink model. The ACM architecture is highly suitable for implementations of SDR systems because of its inherent hardware and software flexibility. In addition to the flexibility obtained by changing software components, the hardware platform offers two-fold flexibility as well. Firstly, as with picoArray, the mapping of the algorithm onto the array of processing elements can change rapidly at run-time. Secondly, and this flexibility is unique to the ACM, each of the processor elements can be rapidly adapted to the processing task at hand.

3.5.2.3.7 Further Developments

Note that there are plenty of other approaches quite similar to the ones reported here. Tricend Corporation offers a combination of standard cores, memory and an interconnect bus, all together named a configurable processing system unit [86]. The University of Washington have developed their RaPiD emulator [87] and Carnegie Mellon University's PipeRench [88] and Brigham Young University's DISC [89] are to also mention.

3.6 Epochs for the Developments of RF Linearisation Technology

Table 3-5 gives an indication of availability for the three different design characteristics identified in Section 3.5.

Technique	University Proposal	First Start-ups	Mature Technology Expected
Fully Fixed		2002	2006 – 2007
Flexible at Design-time	2002	1999	2008 – 2010
Flexible at run-time	1997	2000	2012 – 2015

Table 3-5: Epoch expectations for SDR Technologies related to RF Linearisation

3.7 Conclusions

With flexible radio architectures of the future comes the demand for high bandwidth signals. Since low-power power amplifiers are an indispensable part of the transmission chain, the demand for linearisation methods is very high. When considering the high signal bandwidth requirement, only digital pre-distortion methods currently have the potential to supply such linearisation. Due to their high bandwidth, power amplifiers show not only nonlinear but also dispersive effects that are difficult to handle in a general form. Thus adaptive algorithms are required with parameters and structure that can be changed depending on the current transmission scheme.

With the emerging concept of SDR, new flexible platforms seem to be at hand, capable not only of providing a reprogrammable hardware platform but also of supporting the various needs of an optimal adaptive algorithm for pre-distortion. In particular those hardware architectures that are flexible during run-time are of particular interest for SDR as well as for adaptive pre-distortion techniques. Since such techniques require sophisticated algorithms running permanently to adapt to continually changing behaviour, it is of the utmost importance to have hardware architectures available that are as optimal as possible in order to achieve the best results with the lowest complexity. While first adaptive algorithms dealing with broadband power amplifiers are emerging, only the most flexible hardware architecture can provide the platforms needed to implement them with low power consumption and low complexity. Once such flexibility has been sacrificed, the platform can only be optimal for a specific algorithmic variant, at a specific bandwidth and for a specific amplifier. The concept of SDR will allow highly complex algorithms for pre-distortion to be utilised in base stations as well as in cellular phones.

The emerging concept of SDR enables new flexible platforms, capable of providing reprogrammable hardware and supporting optimal adaptive algorithm for pre-distortion. The optimal algorithm for adaptive RF linearisation is not known. This field is relatively new and much is in progress and in movement. It is very likely that there will be different algorithmic techniques depending on the mode of operation rather than a single one.

A table has been generated which indicates epochs for the development and the maturity for SDR technologies related to RF linearisation.

4 Antenna Processing



By Stephan Weiss, University of Southampton.

4.1 Introduction

This chapter considers antenna processing. It is a fundamental motivation in software defined radio (SDR) technology to move the boundary between analogue and digital domains in both the radio transmitter and receiver closer to the antenna in order to achieve reconfigurability, flexibility, or simply to permit the implementation of advanced digital signal processing algorithms for enhanced transmission and spectral efficiency.

In this chapter we focus on the processing of the antenna signal, i.e., the conversion of an RF signal in the receiver into a digital signal, and the transformation of a digital signal into an analogue RF signal in the transmitter. The crucial elements in this process are the conversion devices. Therefore, following a brief overview of general radio receiver and transmitter architectures that form candidate schemes for SDR systems (Section 4.2), we will discuss the key fundamentals of digital-to-analogue and analogue-to-digital conversion (Section 4.3).

This discussion naturally leads to analogue-to-digital converters (ADCs) and digital-to-analogue converters (DACs). Considering commercial products for conversion devices, the current bottleneck is the ADC. With respect to SDR, excellent summaries on this topic have been published in the past by [65][90] and, in Tuttlebee's edited SDR volume in 2002 [80][91]. In this section we aim to add to these documents an overview of the state-of-the-art in 2005 (Section 4.4), and to review possible future technologies that could enhance the performance of conversion devices for SDR systems (Section 4.5).

4.2 Receiver and Transmitter Architectures

The aim of the transmitter is to convert a complex baseband signal, characterised by an in-phase and quadrature component, into a real valued signal oscillating at a given carrier frequency. In the receiver, the task is reversed; a complex valued baseband signal has to be extracted from an RF signal. In terms of hardware realisation, both transmitter and receiver often utilise the same RF antenna by means of an RF switch (e.g., for time-division duplex (TDD) operation) or a duplexer (e.g., for frequency-division duplex (FDD) operation).

The remaining hardware building blocks between RF and the digital processing stage, albeit in general dual in their functionality, have to be realised separately for transmission and reception. In the following sections we will present some example receiver and transmitter architectures suitable for use in SDR applications.

4.2.1

Receiver

In the receiver, the RF output from the antenna is passed via the duplexer to a low-noise amplifier (LNA). The LNA is an amplifier that is constrained to minimally affect the RF signal in terms of non-linear distortion and noise corruption. Due to these constraints, the LNA generally only applies a moderate gain to the received signal. Since signal levels for various applications and standards can vary considerably [92], adaptive (or automatic) gain control (AGC) is required in SDR systems. This is not applied in the LNA, but in one of the subsequent amplification stages in the signal chain of the receiver. The receive signal chain is reviewed briefly below, since various possibilities and candidate schemes for receiver architectures that are applicable to SDR systems exist.

4.2.1.1 Direct Conversion or Zero IF Structure

In a direct conversion receiver, a bandpass filter is used to select an RF band of interest. This band is quadrature demodulated to baseband (i.e., to an IF of zero, hence the term zero IF) by mixing with a local oscillator (LO). After mixing, the in-phase and quadrature components are converted to the digital domain separately, as shown in Figure 4-1. The input signals to the ADCs are typically pre-processed by an amplifier with large gain and an AGC feature to make optimal use of the dynamic range of the ADC.

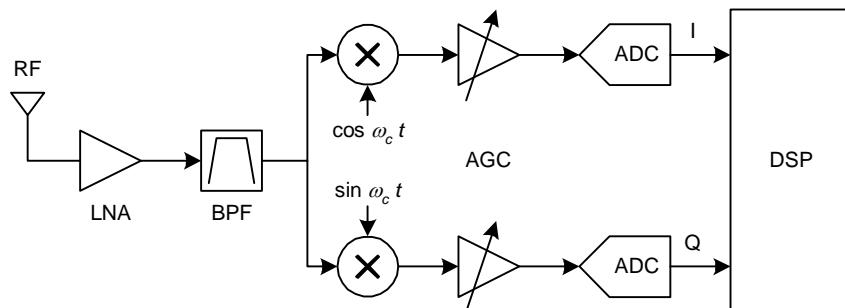


Figure 4-1: Direct conversion of zero IF receiver architecture

The advantages of direct conversion architectures are the simplicity of the filter arrangement, the low complexity, and the convenience of sampling a complex signal (as will be highlighted in Section 4.3.1). The major disadvantages are considered to be the possible introduction of a direct current (DC) offset, the in-band distortion by modulation products, and the realisation of a wideband analogue quadrature demodulator [92].

A DC offset can arise through coupling of the LO into the RF stage, from where it is fed back into the demodulation stage. Self-mixing then produces the DC component; although there are other sources of DC offset, this is considered to be the major contributor. Distortion products occur due to mixing. If two frequency components, e.g., f_1 and f_2 , are demodulated together, then second-order and third-order distortions will appear at, e.g., $f_1 - f_2$ and $2f_1 - f_2$, respectively.

UNCLASSIFIED

In standard superheterodyne receivers, such distortion products are controlled by performing the down conversion in stages, with filters, adjusted to remove unwanted products, between stages. However, since the direct conversion receiver would be anticipated to operate over a considerable frequency range, such distortion products are likely to lie within the band of interest. Finally, the requirement of the quadrature demodulation stage to be balanced over a wide range of frequencies is generally difficult.

4.2.1.2 Low IF Conversion Receiver

In order to circumvent the DC offset and distortion product problems found in a direct conversion receiver, as well as reduce the operational bandwidth of the analogue quadrature demodulator, its structure can be amended by one or more superheterodyne stages prior to quadrature demodulation and analogue-to-digital conversion. Again, sampling would be performed on the complex valued quadrature demodulated signal, retaining the simplicity of the direct conversion receiver while avoiding its main disadvantages. However, a drawback of this so called low IF conversion scheme lies in the requirement of a series of analogue RF/IF filters with high quality and better image rejection than needed in a direct conversion receiver [92].

4.2.1.3 Multiple Conversion Receiver

Multiple conversion receiver architectures rely on a cascade of superheterodyne stages to downconvert the RF signal to IF, involving a number of filter and amplifier stages. Finally, as shown in Figure 4-2, the real valued IF signal is converted to the digital domain, where a quadrature modulation into the baseband can be performed. The main difference to the previous architectures is that the signal passed to the analogue-to-digital conversion stage is real valued. The advantage of the multiple conversion structure is that the quadrature demodulation in the digital domain is very accurate in terms of phase quadrature and the amplitude of the local oscillator. Further, the multiple conversion receiver shares the benefits of the low IF receiver structure with respect to the gain being adjusted over several stages. Thus, the selectivity can be high and the distortion products produced in various demodulation stages can be controlled and suppressed. Drawbacks arise from the high complexity of the receiver, the potential requirement of several LOs and accurate filters in order to achieve image rejection.

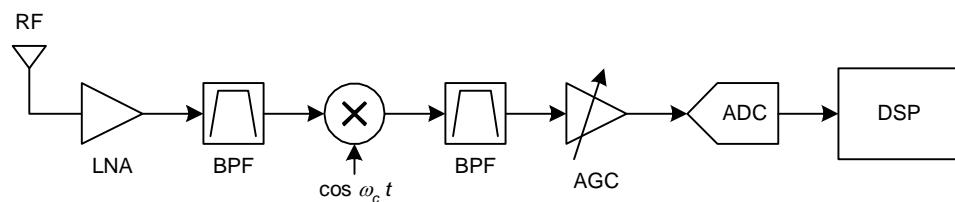


Figure 4-2: Multiple conversion receiver (In-Phase Component)

UNCLASSIFIED

4.2.2 Transmitter

The transmitter reverses the functionality of the receiver, transforming a baseband signal to RF in order to be radiated from the antenna. The RF front-end generally includes bandpass filtering in order to band limit the signal and minimise interference to neighbouring frequency bands. This is followed by a high power amplifier (HPA), prior to passing the signal to the antenna via the duplexer.

4.2.2.1 Direct Conversion Transmitter

A direct conversion transmitter converts a complex signal with in-phase and quadrature components to the analogue domain via two separate DACs. In the analogue domain, the signal is quadrature modulated and passed to the bandpass filter and HPA as described above. The typical structure of a direct conversion transmitter is depicted in Figure 4-3.

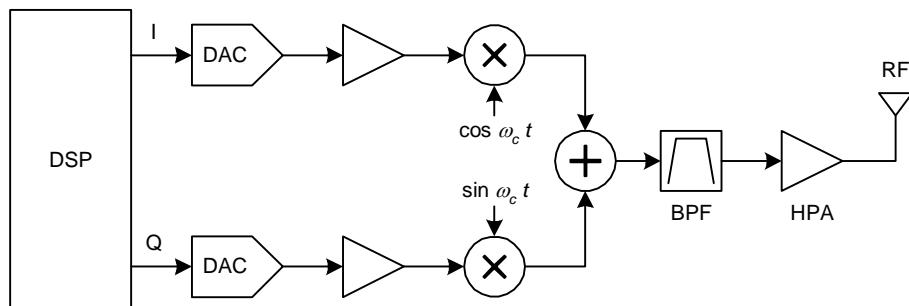


Figure 4-3: Direct-conversion transmitter

As with the direct conversion architecture for the receiver, the advantages of the direct conversion transmitter are the simplicity of the filtering task involved, the low complexity of the method, and the ability to digitally synthesise most of the signal, hence lending superior quality to both signals and filter characteristics. Problems arise due to the bandwidth over which the implementation of the quadrature modulation needs to be balanced and accurate, and over which the HPA and the filters have to provide high quality and linearity.

4.2.2.2 Multiple Conversion Transmitter

A multiple conversion transmitter potentially performs the up-conversion from IF to RF in several stages, some of which may be performed within the digital domain, e.g., a DSP or FPGA. The in-phase and quadrature components of the digital complex signal are separately converted and passed through a reconstruction filter with bandpass characteristic in order to select an appropriate frequency band from the spectral repetitions after digital-to-analogue conversion. A typical architecture is shown in Figure 4-4. Its main distinction from a direct conversion transmitter, such as that shown in Figure 4-3, is the lowpass filtering required for the reconstruction following the DAC, and the potentially one or more superheterodyne stages for up-conversion.

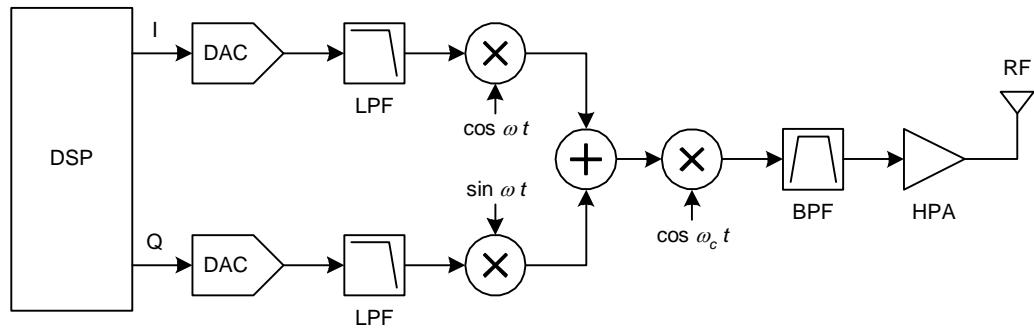


Figure 4-4: Multiple-conversion transmitter

A good background on the various receiver and transmitter architectures, together with a closer look at their performance measures and characteristics, is provided in [80] and [92].

4.3 Analogue-to-Digital and Digital-to-Analogue Conversion

The viability of the receiver and transmitter structures reviewed in Section 4.2 depends on suitable interfaces between the analogue and digital domains. In order to push the boundary between analogue and digital domains towards higher IFs or even beyond, both the ADCs and DACs require a sufficiently high sampling rate. Furthermore, high resolution in terms of the quantiser's word length is important for achieving a large dynamic range. Additionally, low sampling jitter, good linearity and low power consumption and cost are further desirable properties [65][90][91]. In this section we will provide an overview of ADC and DAC theory and the generally perceived requirements for these devices in SDR applications.

4.3.1 Sampling

A first step towards a digital number representation of an analogue signal, $x(t)$, with t being the continuous-time variable, is the sampling process to obtain discrete-time values. This is generally performed by a sample-and-hold device. If we assume ideal sampling of an analogue electrical signal, $x(t)$, with a sampling period T_s (and hence a sampling rate of $f_s = 1/T_s$), the result of the sampling process, $x_s(t)$, can be expressed by multiplying $x(t)$ with a train of pulses spaced at integer multiples of the sampling period. As a result, the spectrum of the sampled signal $x_s(t)$ will be periodic with respect to the sampling rate.

4.3.1.1 Nyquist Sampling

For Nyquist sampling, the signal $x(t)$ is sampled at a rate equivalent to at least twice the maximum frequency component contained in $x(t)$. An example of Nyquist sampling is shown in Figure 4-5. The spectrum of the continuous-time signal, shown in Figure 4-5 (top), becomes periodic as a result of the sampling process, as shown in Figure 4-5 (bottom). If $X(j\omega)$, the Fourier transform of $x(t)$, is sufficiently band limited, the spectral repetitions do not overlap, thus avoiding aliasing.

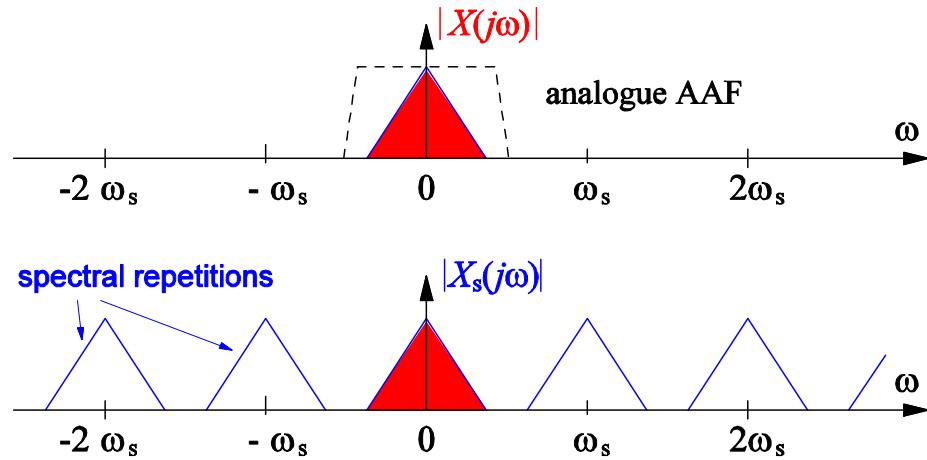


Figure 4-5: Nyquist or baseband sampling: the spectrum of an analogue input with indication of the magnitude response of the anti-aliasing filter (top) and the spectrum after ideal sampling (bottom)

We need to ensure that sampling at the Nyquist rate provides a correct and faithful representation of the analogue signal $x(t)$ in the digital domain, such that the underlying analogue signal can be reconstructed. As the exact spectral content of the signal supplied to an ADC is *a priori* unknown, the signal has to be appropriately band limited to avoid potential aliasing. This is performed by a lowpass anti-aliasing filter (AAF) with a stopband frequency of half the Nyquist rate, as shown in Figure 4-5 (top).

4.3.1.2 Bandpass Sampling

If the analogue signal is not a baseband signal but rather has bandpass characteristics, it is possible to acquire the signal at a rate slower than the Nyquist rate. This process is known as bandpass sampling or undersampling and, in the context of SDR, is often referred to as IF sampling. The theoretical ease of bandpass sampling depends on whether the analogue signal is complex valued - ideally analytic, or not.

For analytic or complex valued bandpass signals, such as shown in Figure 4-6 (top), the minimum sampling rate is given by the signal's bandwidth. The sampling process will make the spectrum periodic, as shown in Figure 4-6 (bottom). If the sampling rate is chosen to be higher than the bandwidth of the signal, no overlap of the spectral repetitions will occur and hence problems due to aliasing are avoided. Therefore, complex valued bandpass sampling is straightforward and only needs to consider the signal bandwidth. Note that the resulting sampled signal remains complex valued, as indicated by the lack of symmetry in the spectrum in Figure 4-6 (bottom).

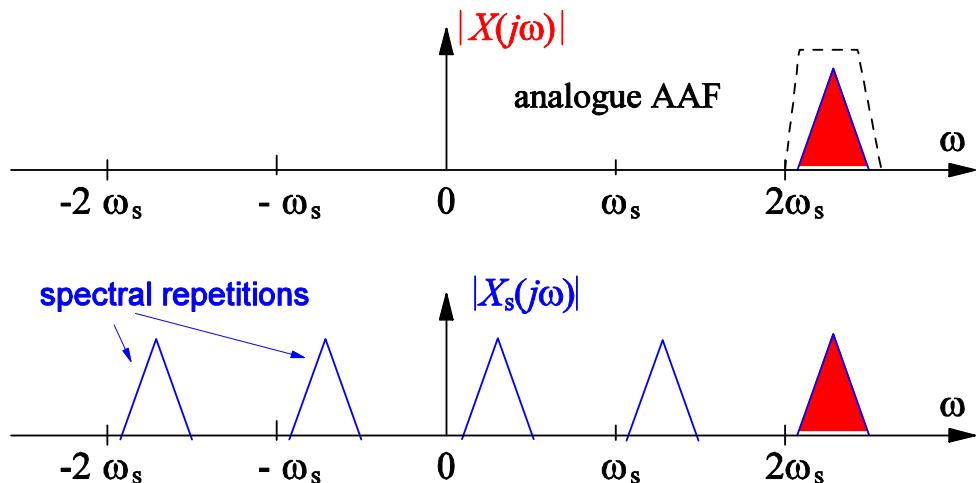


Figure 4-6: Bandpass sampling or undersampling of an analytic signal showing the spectrum of an analytic quadrature demodulated analogue IF input with anti-aliasing filter (top) and the spectrum after undersampling (bottom)

Undersampling a real valued bandpass signal is not straightforward and requires careful consideration of both the signal's bandwidth and its band position. An example is provided in Figure 4-7, where the symmetric spectrum in Figure 4-7 (top) highlights the real valued nature of the analogue input. This real valued signal can be regarded as a superposition of two analytic signal parts, one having non-zero components for positive frequencies only, the other one for negative frequencies only.

Given the linear time-varying property of sampling [93], in the sampling process both analytic signals are periodised separately. In the example in Figure 4-7, the superposition of both periodised spectra leads to no overlap and hence no aliasing. However, more planning is required compared to the complex valued case, and strict rules on the bandwidth and band position have to be obeyed in order to not incur any aliasing [94][95][96].

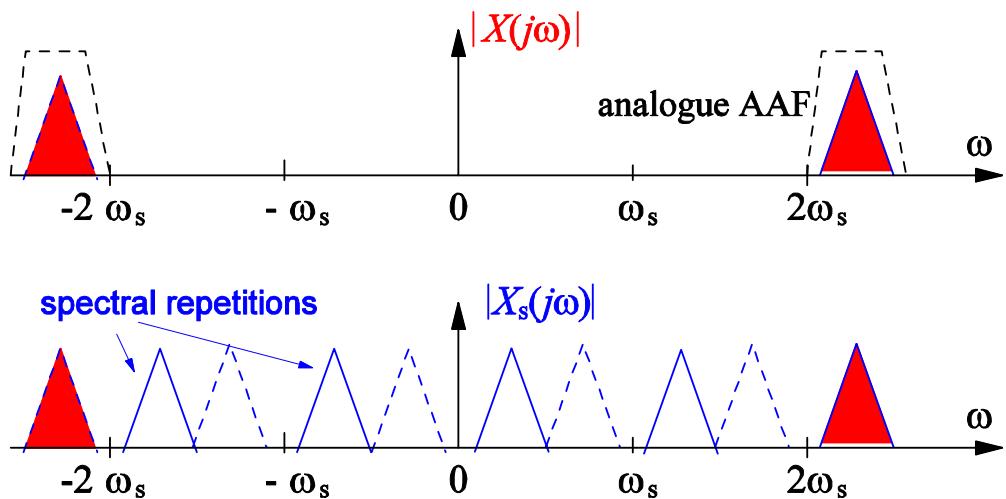


Figure 4-7: Bandpass sampling or undersampling of a real-valued signal showing the spectrum of a real-valued signal with anti-aliasing filter (top) and the spectrum after undersampling (bottom)

While bandpass sampling of real valued signals can be difficult, the sampling of analytical signals is straightforward [96][97], as explained earlier. Quadrature demodulated signals fall into the later class, whereby in-phase and quadrature component have to be sampled separately, thus requiring two ADCs. Different from Nyquist sampling, the analogue anti-alias filter for bandpass sampling needs to have a bandpass characteristic as indicated in Figure 4-6 and Figure 4-7.

4.3.1.3 Implementation

Sampling is mostly performed by sample-and-hold devices implemented as switched-capacitor circuits, whereby, for a short aperture time, a capacitor is charged according to the current voltage value of the analogue waveform to be sampled. Thereafter, the capacitor value is passed on to a quantiser.

UNCLASSIFIED

Errors in sampling occur in various forms. Timing jitter refers to uncertainties in the timing of the sampling clock, describing sample-by-sample variations by their root mean square (RMS) interval of uncertainty. Timing jitter is sometimes also referred to as aperture uncertainty. In contrast, aperture errors refer to the deviation from an ideal pulse of the opening time during which the capacitor is exposed to the incoming waveform. Aperture errors arise from transition periods of the switch between ‘open’ and ‘closed’ states with a finite rise and fall time, as shown in Figure 4-8. This leads to a sampled signal that is obtained by multiplying the analogue waveform by a train of non-ideal pulses rather than the pulse train assumed for ideal sampling. If aperture errors are significant, the sampled signal will suffer from spurious non-linear components [91].

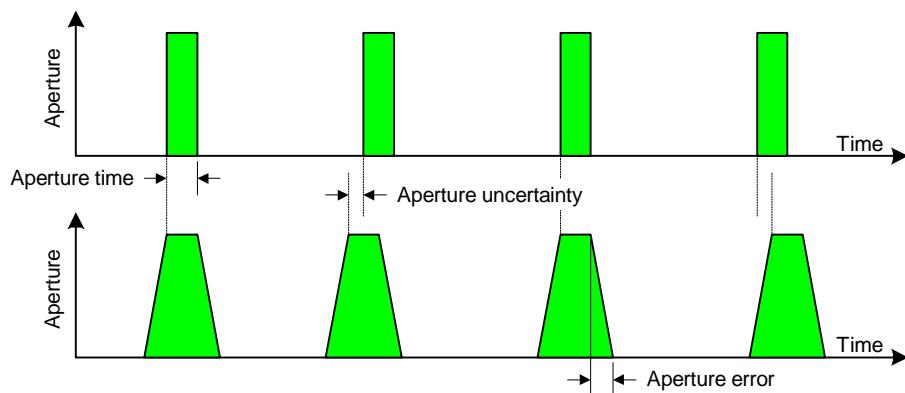


Figure 4-8: Ideal aperture for sampling (top) and aperture uncertainty and aperture error (bottom)

4.3.2 Quantisation

A quantiser converts the continuous voltage value of the sample-and-hold output to a number that can be represented by a given binary word length R . This process is not unambiguous and generally voltage values within a range q will be rounded to the same level representing a specific word. An example is provided in Figure 4-9, where a quantiser with 3-bit number representation codes $2^3 = 8$ discrete levels, to which the quantiser input is mapped by means of a quantiser characteristic.

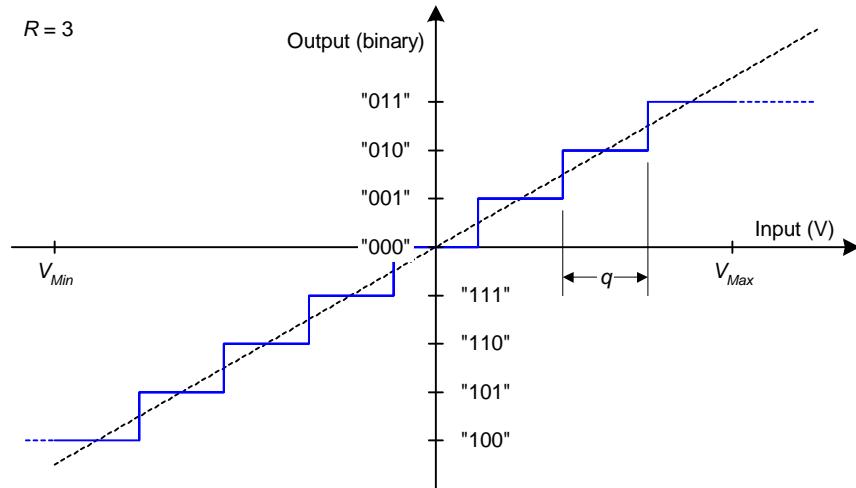


Figure 4-9: Quantiser characteristic

Mathematically, a quantiser is modelled as adding quantisation noise to the continuous valued input samples of the quantiser. Therefore, a popular measure for the quality of the quantiser is the signal-to-quantisation noise ratio (SQNR). For uniformly distributed quantisation noise, the quantisation noise power is given by $q^2 / 12$; if the input signal excites the full input voltage range of the quantiser characteristic, then the SQNR can be calculated as [98]:

$$\text{SQNR} = 6.02 \times R + 1.76 \text{ dB} \quad \text{Equation 4-1}$$

If the quantiser characteristic approaches a straight line through the origin, as given in Figure 4-9, the quantiser characteristic is referred to as linear. Non-linear characteristics, where a different resolution is provided at different voltage ranges, can give a benefit in terms of SQNR, if the characteristic is optimised according to the probability density function of the analogue input signal. Examples of such characteristics are A-law and μ -law quantisers, which improve the SQNR for a speech input from eight bits to virtual resolutions of 12 - 13 bits. For SDR applications, linearity is generally a desirable property [91], and deviations through element inaccuracies within the circuit result in nonlinearities, causing harmonic distortion and the emergence of spurious components in the signal spectrum. The literature distinguishes between DNL and INL, depending on which axis (i.e., the input voltage range or the levels at the output) suffers from non-uniform spacing.

In an ADC, quantisation noise is not the only form of distortion imposed on the signal. For example, thermal noise is considered as the limiting factor for ADCs with very high resolution [90]. Therefore, in addition to the number of bits utilised, manufacturers will in general state an effective number of bits (ENOB), giving the ratio between a signal power just within the quantiser range and the quantiser's noise floor evaluated from device measurements. ENOB is calculated by rearranging the above equation for SQNR:

$$\text{ENOB} = \frac{1}{6.02} \times \left[10 \times \log_{10} \left(\frac{\text{signal power}}{\text{power of measured noise floor}} \right) - 1.76 \right] \text{ bits} \quad \text{Equation 4-2}$$

For most manufactured ADCs, the number of bits deployed in the quantiser characteristic, R , is nearly identical to their ENOB value. For experimental devices reviewed in Section 4.5, however, ENOB is an important measure and provides a fairer assessment than SQNR according to the above formula.

4.3.3 Quality Measures

To assess the performance of an analogue-to-digital converter, Sections 4.3.1 and 4.3.2 have introduced the sampling rate, measured in samples per second (Sps), and ENOB, measured in bits. Further, sampling jitter and aperture errors have been defined as important characteristics limiting the performance of an ADC, whereby jitter can be assessed by the RMS time interval of uncertainty. Beyond these discussed effects and measures, a number of further quality criteria exist. While most of them are less significant [91], the SFDR is important in connection with the receiver chain. SFDR evaluates the difference in power between the noise floor and the power of a signal that would just cause a non-linear distortion component to emerge from the noise floor [92]. One such cause can be the aperture error discussed in Section 4.3.1.

4.3.4 Requirements

Based on the performance characteristics of ADCs reviewed in the previous sections, some comments on the requirements of data converters, including DACs, can be made with respect to their utilisation in SDRs. The requirements are dependent on whether conversion devices are deployed in a base station or handset [91], and in the receive or transmit part of a radio link.

4.3.4.1 Base Station vs Mobile Handset

Handsets impose restrictions on the size of devices as well as their power consumption in order to permit small casing and a long battery life. Further, the number of mobile handsets and competitiveness of the market favours low cost solutions. Therefore, factors such as size, power consumption, as well as the commercial cost of conversion devices are decisive in the handset.

UNCLASSIFIED

At the base station it can be desirable to convert multi-carrier signals and hence replace multiple radios, each operating on one or a subset of carriers, by a single front-end device [91]. Hence ADCs and DACs in the base station may require larger bandwidths, which can here be traded off against the size, power consumption and cost of the conversion devices. However, the fact that environmental noise restrictions apply to base stations, particularly when deployed within residential areas, means that the only cooling is often by heat dissipation rather than by cooling fans. Therefore, some operators, such as Orange, consider power consumption in the region of several Watts for a conversion device too high.

4.3.4.2 Transmitter vs Receiver

In the receiver, irrespective of deployment in a base station or mobile handset, in the past, 14 bits of resolution have been considered necessary to be able to receive weak signals in the presence of stronger ones [91][99]. The difference in power between different signals may have different origins. Firstly, within a frequency band with a multiple access scheme, the near-far effect may cause some user signals to be significantly weaker than others. Secondly, if standards are to be integrated, different power sensitivities apply to different transmission regimes [92] and must be considered in the receiver design. While ADCs with 14 bit resolution have been utilised for most SDR test beds that permit a limited standard integration [100][101][102], future SDRs that can manage a whole host of standard implementations might require higher bit resolutions due to the potentially strong difference in power levels defined for the candidate schemes to be combined in a single device.

For transmitters, the DAC is, in general, not as critical since commercially available devices are generally faster, cheaper and have higher resolution than their commercially available ADC counterparts. With respect to a required resolution, currently a word length of 12 bits is applied in most SDR prototypes [100][101][102] and is often deemed sufficient.

4.4 Commercially Applied Technology and Available Devices

Based on the requirements derived in the previous sections for conversion devices in SDR systems, a number of conversion architectures currently available on the market are reviewed in the following sections. By comparing their performance with past predictions for the performance of conversion devices, we aim to extrapolate the evolution of ADC and DAC technology for the near future. Before discussing state-of-the-art ADCs and DACs, important technologies and architectures for ADC and DAC devices are reviewed.

4.4.1 ADC and DAC Conversion Techniques

4.4.1.1 General ADC Components

Analogue-to-digital conversion is based on preconditioning, sampling and quantisation of an analogue waveform, as reviewed in Section 4.3. The preconditioning stage consists of an amplifier controlling the gain of the incoming signal such that the input voltage range of the ADC's quantiser characteristic is utilised as best as possible. The amplifier also needs to be adjusted such that the quantiser range is not exceeded, which otherwise results in clipping and non-linear, harmonic distortion. The preconditioning stage is further comprised of an analogue anti-alias filter of appropriate form according to Section 4.3.1.

The essential device for sampling is a sample-and-hold circuit, also often referred to as track-and-hold, which is currently based on switched capacitor techniques [90]. In current CMOS technology, the sampling process is subject to jitter and a non-ideal aperture, with transitions between the open and closed states, resulting in undesired effects as outlined in Section 4.3.1. Jitter is considered a severe problem and a technology limitation in CMOS devices at high sampling rates [91]. Bandpass sampling at high rates requires very short and accurate aperture times since, with respect to the sampling interval, the input signal varies considerably faster than in Nyquist sampling.

Although a large number of different techniques exist for quantisation of the acquired samples, such as ramp or successive approximation, the high specifications imposed by SDR with respect to both resolution and sampling rate generally limit the choice to pipelined Flash and sigma-delta converters [91][99]. These techniques are reviewed below.

4.4.1.2 Flash ADC

A Flash or direct conversion ADC with R bits resolution contains 2^R comparators, which compare the analogue input - passed on from a sample-and-hold circuit - against the outputs of a resistor ladder providing reference voltages. As shown in Figure 4-10, the result of this comparison is held in latches and converted to a binary number by a look-up table [103] or a so-called thermometer [91].

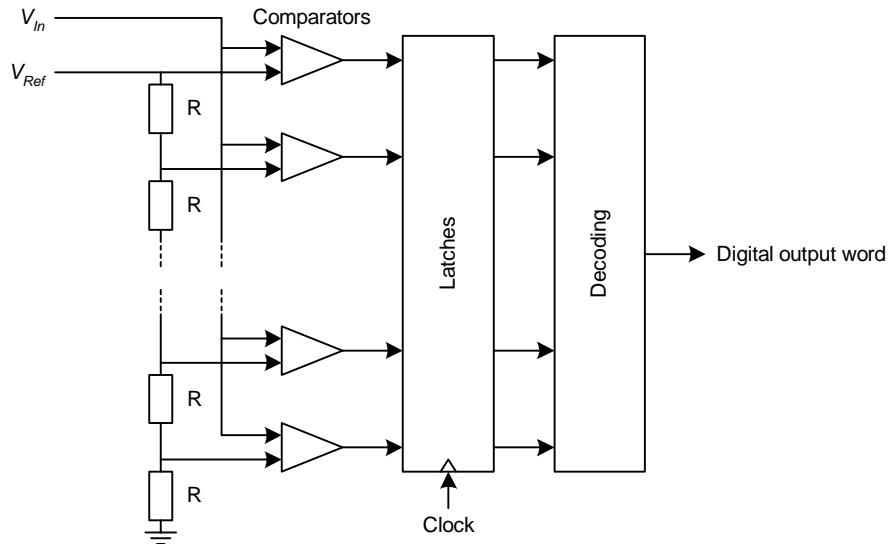


Figure 4-10: Flash ADC

A Flash ADC provides a very fast conversion, but potentially requires a very large number of comparators for longer quantisation word lengths. Therefore, in practice, there are few Flash ADCs with a word length of more than 10 bits [91]. The linearity of the device depends on a very low tolerance of the elements in the resistor ladder. Further complications are comparator ambiguity, which becomes evident at very high sampling rates, and is seen as a performance bound when sampling beyond several tens of gigahertz [90].

4.4.1.3 Pipelined or Multi-Stage ADC

A pipelined ADC uses a cascade of fast Flash conversion sections. By cascading, a higher resolution can be achieved compared to standard Flash devices, whilst containing the complexity of the ADC. The cascade consists of a number of identical stages comprising switched-capacitor circuits for sampling and a low resolution Flash ADC section. When the input signal is applied, each stage samples and quantises the signal to a low resolution and subtracts the quantised voltage value at the stage output from the stage input. The residue between the two is then amplified and passed on to the next stage. The amplification ensures that the same implementation can be employed for subsequent stages, while the quantiser resolution becomes finer as the residues progress down the cascaded architecture. Most ADCs that are suitable, and utilised, for SDR applications are based on such a pipelined Flash architecture.

UNCLASSIFIED

4.4.1.4 Sigma-Delta Converters

Sigma-delta conversion is based on the principles of oversampling and noise shaping [104][105]. Modelling the quantiser in an ADC as adding white Gaussian quantisation noise to the signal of interest, the quantisation noise power only depends on the word length R , and is spectrally flat. By oversampling the ADC, the signal of interest only occupies a limited bandwidth, while the quantisation noise is spread over the entire spectrum. By applying a digital filter matched to the signal bandwidth to the quantiser signal, out-of-band quantisation noise is suppressed. This suppression enhances the signal-to-quantisation noise ratio. Every quadrupling of the oversampling factor is equivalent to gaining an extra bit in resolution.

In addition to oversampling, noise shaping is employed in the sigma-delta architecture, whereby the analogue signal is filtered, usually by an analogue filter prior to sampling with r bits of resolution at an oversampled rate. The r -bit output is converted to a voltage signal by an r -bit DAC and negatively fed back to the analogue input of the device. The characteristics of the circuit are such that the signal of interest does not experience any frequency dependent gain, while the power spectral density (PSD) of noise is characterised by the inverse of the analogue filter. Therefore, if the filter is lowpass, the quantisation noise PSD will have a highpass characteristic. In the subsequent digital filtering and decimation stages, the noise suppression known from oversampled devices is enhanced due to the fact that the quantisation noise power is attenuated in band by the inverse filter characteristic.

Sigma-delta devices can perform bandpass sampling if the analogue filter inside the device has an appropriately selected bandpass characteristic. High-performance sigma-delta converters can be achieved by selecting $r > 1$, through appropriate quality of the analogue filter, and through the use of higher order feedback architectures, although stability can become an issue [105].

4.4.1.5 Digital-to-Analogue Converters

High-speed DACs are generally CMOS based and perform the reconstruction of a voltage signal through the summation of resistor-weighted currents. For the sampling rates and bit resolutions required in SDRs, manufactured DACs often rely on a segmented multi-stage arrangement [91] of the current switches.

Although possible in theory [93], DACs are not undersampled but are operated at least at the Nyquist rate. The voltage output of the DAC contains steps, which results in repetitions of the baseband spectrum. Analogue interpolation or reconstruction filters suppress these image spectra. The quality of these analogue reconstruction filters can be traded off against additional processing in the digital domain. For example, if the DAC is operated in an oversampled mode and the upsampling and interpolation filtering is performed in the digital domain prior to conversion, the position of the first image is shifted to a higher frequency, thereby relaxing the requirements of the analogue reconstruction filter. An example is shown in Figure 4-11.

UNCLASSIFIED

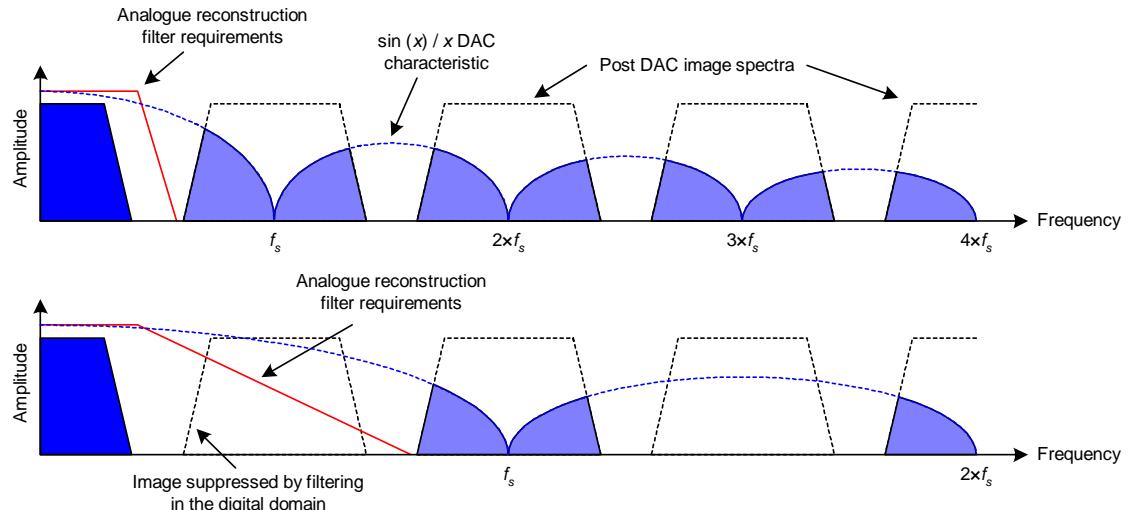


Figure 4-11: Example showing the use of digital signal processing and a two-times oversampling DAC to relax the requirements of the post-DAC analogue reconstruction filter (bottom) compared to a DAC operating just above the Nyquist rate (top)

4.4.2 Commercial ADC and DAC Devices

In the following sections we will summarise the performance characteristics of a number of commercially available, state-of-the-art ADCs and DACs. The main manufacturers of converters are Analog Devices (AD) and Texas Instruments (TI), on whose products Table 4-1 (for ADCs) and Table 4-2 (for DACs) focus.

4.4.2.1 Analogue-to-Digital Converters

The ADCs listed in Table 4-1 give an idea of the maximum sampling rates that can currently be obtained at various word lengths. At 16-bit resolution, the TI ADS1605 offers a reasonable cost solution for a bandwidth that is rather narrow in SDR terms. AD's AD10678 offers a good bandwidth but at the expense of a high power consumption and high purchase cost. In the 14-bit range, which is considered necessary in order to permit a sufficient dynamic range in a receiver, the fastest converter currently available is TI's ADS5500, which has a sampling rate of 125 MHz. Substantially higher sampling rates are available at 12-bit resolution with the AD12400 and the AD12500, which is soon to enter the market. The latter devices are based on a technique that uses two ADCs, time-interleaved to achieve a higher sampling rate. We will consider this technique in greater detail in Section 4.5.1. The downside of these fast sampling devices however is their fairly high power consumption, which renders them unsuitable for use in handsets and may even prevent their deployment in base stations, considering the comments made in Section 4.3.4. At eight bits, Dallas Semiconductor Corp./Maxim Integrated Products Inc. offer a Flash converter with a maximum sampling rate of 1.5 GHz.

UNCLASSIFIED

Man.	Component	Rate	Resolution	Power	Cost	Comment
TI	ADS1605	5 MSps	16 bit	0.56 W	\$33	Sigma-delta
AD	AD10678	80 MSps	16 bit	8 W	\$580	
AD	AD6645-105	105 MSps	14 bit	1.5 W	\$88	Bandpass, 200MHz
TI	ADS5500	125 MSps	14 bit	0.78 W	\$95	Pipeline
TI	ADS5541	105 MSps	14 bit	0.71 W		Pipeline
AD	AD12500	500 MSps	12 bit			Released 2005
AD	AD12400	400 MSps	12 bit	8.5 W	\$450	Time-interleaved, Nyquist only
AD	AD9433-125	125 MSps	12 bit	1.5 W	\$76	Bandpass, 150MHz
TI	ADS5273	70 MSps	12 bit	1.1 W		Pipeline
AD	AD6640	65 MSps	12 bit			Bandpass, 300MHz
TI	ADS5221	65 MSps	12 bit	0.285 W	\$14	
Maxim	MAX108	1.5 GSps	8 bit			Flash

Table 4-1: Commercial ADCs aimed at SDR applications

4.4.2.2 Digital-to-Analogue Converters

A range of currently available high performance DACs that are potentially applicable to SDR implementations are listed in Table 4-2. Compared to the characteristics of their ADC counterparts, DACs are, in general, available with higher performance specification at a lower cost and lower power consumption. However, commercial DACs are only manufactured for baseband reconstruction, i.e., an equivalent to bandpass sampling in the ADC stage is not possible in the digital-to-analogue conversion stage.

Man.	Component	Rate	Resolution	Power	Cost
AD	AD9726	600 MSps	16 bit		
TI	DAC5686	500 MSps	16 bit	0.44 W	\$38
AD	AD9777	400 MSps	16 bit	0.41 W	\$40
AD	AD9736	1.2 GSps	14 bit	0.38 W	
TI	DAC5674	400 MSps	14 bit		\$25

Table 4-2: Commercially available DACs

4.5

Research Development and Future Devices

Given the advanced performance for DACs over their ADC counterparts, as noted in Section 4.4, the bottleneck is imposed by the performance limitations of ADC devices available to date. Current research is guided by attempts to overcome or bypass such limitations. Walden [90] provides an excellent overview of the various constraining problems that are imposed on the possible range of sampling rates and bit resolutions that can be achieved, as also highlighted in Figure 4-12:

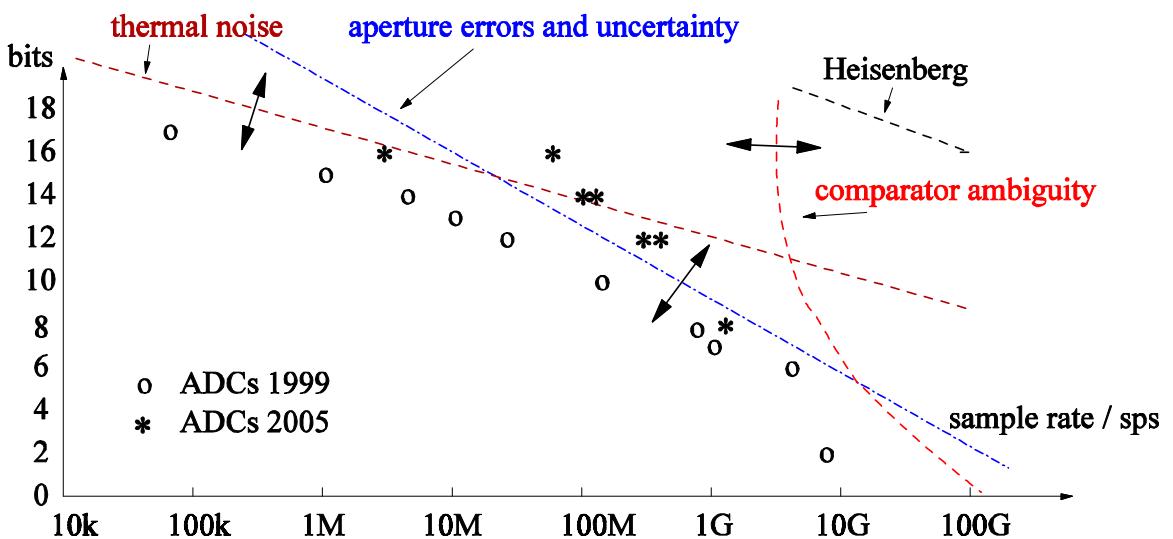


Figure 4-12: Performance of commercially available ADCs

- At low sampling rates, thermal noise is the limiting factor. Depending on the scenario, a drop from approximately 22 bits at 10 kHz to 19 bits at 1 MHz has been stated
- In the mid-frequency range of Figure 4-12, the aperture of the sample-and-hold device limits the performance to a range of approximately 19 bits at 1 MHz to 4 bits at 10 GHz
- Ambiguities in the comparator (for Si CMOS) set limitations for the maximum achievable sampling rate to approximately 10 GHz, even at very low bit resolutions.

Some directions that may overcome or bypass the above limitations have been discussed in [64][99][106]. Research-based developments of ADC devices include sigma delta ADCs [64], pipeline and Flash architectures [64][67][99], as well as cryogenic superconducting devices [106]. In addition, we below focus on time-interleaved architectures [90][107] and recent findings for optical sample-and-hold circuits.

4.5.1 Time-Interleaved ADCs

In time-interleaved analogue-to-digital conversion architectures, several ADC devices are operated in parallel, whereby their sampling clocks are carefully offset against each other [107][108]. An example is shown in Figure 4-13 for two ADCs sampling a waveform in an alternating fashion. The output samples of the various ADCs are then time multiplexed to obtain a signal over the intended bandwidth, also known as a polyphase structure [93]. Recently, this technology has translated into commercial products. For example, the AD12400 listed in Table 4-1 comprises of two time-interleaved ADCs.

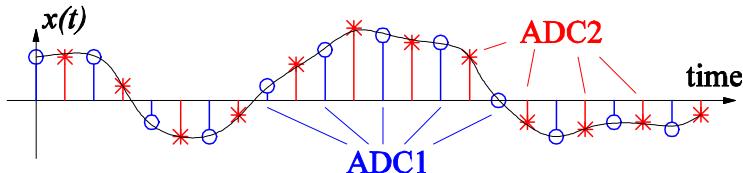


Figure 4-13: Two ADCs in a time-interleaved architecture sampling a signal

While high bandwidth can be achieved with this technique, the drawbacks are in cost and power consumption, as evident in the example of the AD12400 in Section 4.4.2. Technological challenges arise due to misalignment of the various sampling clocks, resulting in non-uniform sampling [95], for which solutions through digital post-processing of the acquired signal have been suggested [109]. Using such judicious post-processing, the number of time-interleaved ADCs could be raised without sacrificing performance due to misalignment of the sampling clocks. Applying this idea, the major remaining limitation lies in the aperture error and uncertainty of the individual devices.

4.5.2 Optical Sampling

Optically sampled analogue-to-digital converters combine optical sampling with electronic quantisation to enhance the performance of electronic ADCs [110]. This technology relies on the possibility of producing very short and very accurate pulsed lasers, which replace the limiting Si-based switched capacitor circuitry. The optical pulses are modulated by the incoming analogue waveform to be sampled. The result is a sequence of pulses whose intensity is proportional to the voltage value of the input. Additional benefits of this technology arise from the availability of optical multiplexing techniques, such that optical sampling can be conveniently combined with time interleaving.

Using low-jitter short-pulse lasers, optically triggered sampling of analogue signals has been reported at 10 GHz [111] and beyond [112]. An experimental system has laid claims to their ability of sampling a 20 GHz signal at a rate of 120 GSps [113], whereby the utilisation in high frequency oscilloscopes has been investigated.

Combinations with other techniques are possible. Bandpass sampling at a sampling rate of 160 MHz at carrier frequencies of up to 40 GHz has been demonstrated in [114]. Similarly, time interleaving has been incorporated in a prototype system [111].

UNCLASSIFIED

While optical sampling may push the boundary imposed by the aperture in purely Si CMOS devices, implementations reported in the literature suggest that the technique might only be applicable to high sampling rates with fairly low resolutions, of the order of 3-4 bits [111][112][114]. Dissenting voices exist, however. In [110], 8.2 bits of effective resolution have been achieved in a system operating at 505 MHz. Moreover, Reference [110] projects that 12 bit resolution in the multi-gigahertz range is possible.

4.5.3 Superconducting Devices

A new family of semiconductor devices is based on rapid single flux quantum (RSFQ) digital logic, exploiting a quantum effect in a superconducting material. Besides a number of other applications, such as sensors and fast logic, this has spawned a new type of ADC and DAC [106][115][116]. Such devices are characterised by a very high-speed acquisition and good resolution. While systems sampling at 100 MHz with an effective resolution of 9.4 bits were operational in 2002 [106], recent developments report sampling at 20 GHz [116].

While literature portrays RSFQs as devices with a very high potential for SDRs [80], the size (and likely cost) of the cooling mechanism required to hold the superconductor at its operational temperature is considerable [106]. However, Reference [106] points out that a number of base stations in the US use cooling systems for RF analogue filters on site, suggesting that RSFQ ADCs and DACs might be deployed in such base stations.

4.6 Epochs for the Developments of RF Antenna Processing Technology

Epoch expectations for ADCs

For ADCs with resolution 6-8 bits we expect a doubling of sampling rate about every 3 years, with the expectation of lower confidence on this broad rule as time progresses. Therefore, this gives rise to decrease in confidence for an estimate occurring at a later period to one sooner (this calculation was undertaken at the end of 2005); Table 4-3 provides the epoch expectations for sampling rate for ADC with 6-8 bit resolution.

ADC with resolution 6-8 bits	
Year	Sampling Rate (GHz)
2001	1
2005	3
2010	6 – 10
2015	12 – 30

Table 4-3: Epoch expectations for sampling rate for ADC with 6-8 bit resolution

UNCLASSIFIED

Epoch expectations for CMOS Power Consumption

The power consumption based on CMOS technology doubles about every two years, again with an assumed increase in error (i.e. a decrease in confidence of the estimate) for years far greater into the future; Table 4-4 provides some predictions.

CMOS Power Consumption	
Year	Power Consumption (mW)
2001	300
2005	1,200
2010	2,000 – 10,000
2015	2,000 – 50,000

Table 4-4: Epoch expectations for CMOS Power Consumption

The CMOS power consumption for 2010 and 2015 can be seen to be very significant and it is possible that the higher power estimated above would CMOS devices impractical in normal usage unless entirely new techniques are being developed by then.

Epoch expectations for DACs

For DACs, higher sampling rates even with higher resolution of 8-10bits can be expected to double every two years. Power consumption is not as critical as in ADCs due to CMOS technology.

4.7

Conclusions

The signal path, from the output of the antenna up to the point of digitisation (and vice versa) has been considered in this chapter. The performance of current ADCs and DACs has been summarised. Most development has been, and remains, targeted at the enhancement of ADC circuits.

In the near future, hybrid structures such as time-interleaved ADCs and the substitution of Si by SiGe technology can be expected to further boost the performance of conversion devices. While such structures are currently often power-hungry, this may not prohibit their deployment in base stations, for example. If novel technologies such as optical sampling and RSFQ techniques mature and perhaps translate into commercially available ADCs, a considerable performance leap into the gigahertz range is expected.

Several tables are provided indicating expected epochs for the development of the technologies associated with antenna processing. Included within this discussion are the epoch expectations for sampling rate for ADC with 6 - 8 bit resolution and CMOS power consumption epoch expectations, which tend to suggest a future need for new technologies for such devices.

5 MIMO Technology and SDR

By Alister Burr, University of York.

5.1 Introduction

Multiple-input, multiple-output (MIMO) techniques for wireless channels, involving the use of multi-element antenna arrays at both the transmitter and receiver, have been identified in the past few years as a means of dramatically increasing the capacity of wireless communication systems and networks. In this section we explore the implications of MIMO for SDR systems. This will include a discussion of both the potential barriers it presents and the complementarities and advantages of implementing MIMO in a software defined radio (SDR) system. We first introduce the principles of MIMO, including the basic theory regarding the capacity and diversity advantages and the main schemes for their implementation, including history and motivation.

Here, we also describe the channel model we will use for realistic assessment of the performance of MIMO systems. We next consider some background literature, firstly exploring previous work on MIMO and SDR and secondly considering the MIMO techniques currently being introduced into the main wireless standards, which are likely to motivate the use of MIMO in any SDR system in the medium term. To determine the potential obstacles to implementation we then consider the implementation of MIMO systems, both in terms of the RF subsystems and of the baseband signal processing, describing alternative architectures at a block diagram level.

Finally, we will consider the advantages of adaptive MIMO schemes that may be regarded as complementary to SDR, and extract some conclusions from the study.

5.2 Introduction to MIMO Techniques

The potential to provide dramatic increases in capacity and spectrum efficiency of wireless systems through the use of multiple antenna arrays at both transmitter and receiver was first pointed out by Foschini [117] at Lucent Bell Laboratories, and subsequently further explored in [118]. It was publicised by a press release from Bell Laboratories in 1998 [119], which raised a great deal of interest in the wireless communications community by appearing to suggest that Shannon's bound on channel capacity [120] had been exceeded. Similar work was also carried out at around the same time at AT&T [121].

Techniques exploiting this effect were termed MIMO (multiple-input, multiple-output). Foschini in [117] proposed a scheme called D BLAST to approach this potential capacity. Subsequently, the same team at Bell Labs developed a more practical scheme, V BLAST [122]. Approaches like these are today usually described by the generic term spatial multiplexing. Tarokh and others at AT&T meanwhile developed a group of schemes called space time codes: space-time trellis codes (STTC) [123] and space-time block codes (STBC) [124]. Their focus, however, was more on achieving the potential diversity improvements available from the use of multiple antenna elements.

UNCLASSIFIED

The fundamental basis of MIMO techniques is to exploit multi-path propagation in the radio channel. Multi-path is an effect (that has traditionally been regarded as deleterious rather than advantageous), which arises when the radio signal travels from transmitter to receiver via multiple paths rather than a single, dominant line of sight path. The multiple paths (or simply multi-paths) occur due to reflection and scattering from objects such as buildings, trees and the general geographic features (or indoors from wall, furniture, etc.), as shown in Figure 5-1. The paths interfere at the receiver to cause Rayleigh fading; if the multiple antenna elements are sufficiently separated, the fading at different elements may be largely uncorrelated, allowing diversity reception.

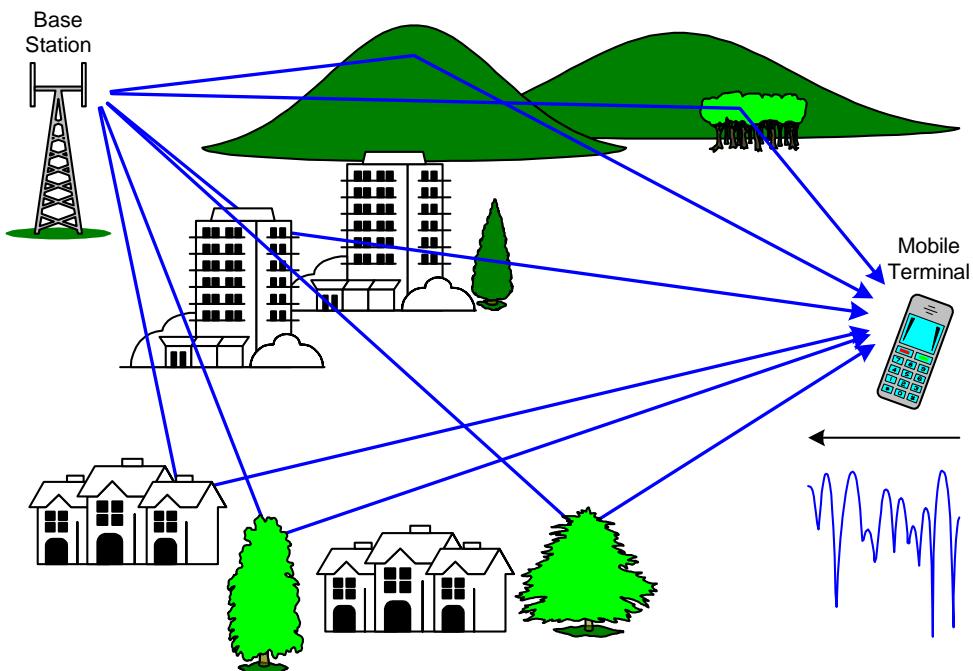


Figure 5-1: Illustration of multi-path propagation

The MIMO channel can conveniently be modelled as an $(n_R \times n_T)$ matrix, universally written as \mathbf{H} , in which the element H_{ik} represents the propagation from transmit antenna k to receive antenna i . We may then write:

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n} \quad \text{Equation 5-1}$$

where \mathbf{r} is a size n_R column vector of (normally complex) signals received on the receive elements, \mathbf{s} is a size n_T column vector of signals transmitted on the transmit elements and \mathbf{n} is a size n_R column vector of noise associated with each receive antenna. n_R and n_T represent the number of receive and transmit elements, respectively. Using this representation the capacity C in bits/s of the MIMO channel was determined in [118], [121] as:

UNCLASSIFIED

$$C = W \sum_{i=1}^n \log_2 \left(1 + \frac{S\lambda_i}{n_T N} \right) = W \log_2 \left(\det \left(1 + \frac{S}{n_T N} \mathbf{H} \mathbf{H}^H \right) \right) \quad \text{Equation 5-2}$$

where W is the channel bandwidth, S/N is the ratio of the total transmit power to noise power per receive antenna and λ_i ($i = 1 \dots n$) are the eigenvalues of $\mathbf{H} \mathbf{H}^H$. $n \leq \min(n_R, n_T)$ is the rank of \mathbf{H} and superscript H denotes the conjugate transpose of a matrix. Note that this assumes that the channel is unknown at the transmitter.

This formula arises from the *singular value decomposition* of the channel matrix \mathbf{H} :

$$\mathbf{H} = \mathbf{V} \Lambda \mathbf{U}^H \quad \text{Equation 5-3}$$

where \mathbf{V} and \mathbf{U} are unitary matrices (that is, their columns all have unit total squared magnitude and are orthogonal to one another) and Λ is a diagonal matrix whose diagonal elements are the square roots of the eigenvalues of $\mathbf{H} \mathbf{H}^H$.

Then if we apply the transformations $\mathbf{r}' = \mathbf{V}^H \mathbf{r}$ and $\mathbf{s}' = \mathbf{U} \mathbf{s}$, we can rewrite Equation 5-1 as:

$$\mathbf{r}' = \mathbf{V}^H (\mathbf{H}(\mathbf{U} \mathbf{s}') + \mathbf{n}) = \mathbf{V}^H \mathbf{V} \Lambda \mathbf{U}^H \mathbf{U} \mathbf{s}' + \mathbf{V}^H \mathbf{n} = \Lambda \mathbf{s}' + \mathbf{n}' \quad \text{Equation 5-4}$$

where $\mathbf{n}' = \mathbf{V}^H \mathbf{n}$. Note that this transformation does not affect the statistics of this noise vector. Note also that the product of a unitary matrix with its transpose conjugate is the identity matrix.

Thus by applying the transformations \mathbf{U} at the transmitter and \mathbf{V}^H at the receiver, we have converted the channel into a set of completely independent channels, which we can use to transmit independent data. These sub-channels can be identified with the *eigenmodes* of the channel. It is the application of Shannon's capacity theorem [120] to these orthogonal sub-channels that leads to the capacity formula (Equation 5-2) – so we note that far from breaching Shannon's bound, the derivation of MIMO capacity, in fact, depends on it.

Note that in the transformations applied above, the columns of \mathbf{V} and \mathbf{U} can be regarded as *steering vectors* applied to the transmit and receive antenna arrays. Weighting and phase shifting the transmitted data according to these vectors forms a particular transmitted beam pattern. Similarly, applying weights at the receiver forms a receive beam pattern. This allows us to visualise spatially how MIMO systems achieve their remarkable capacities, as shown (albeit in a somewhat simplified form) in Figure 5-2. Each column creates a beam pattern in a different direction which is, in turn, carried by a different set of multi-paths. Hence the data are transmitted over channels which are separated spatial, but may reuse the same frequency band.

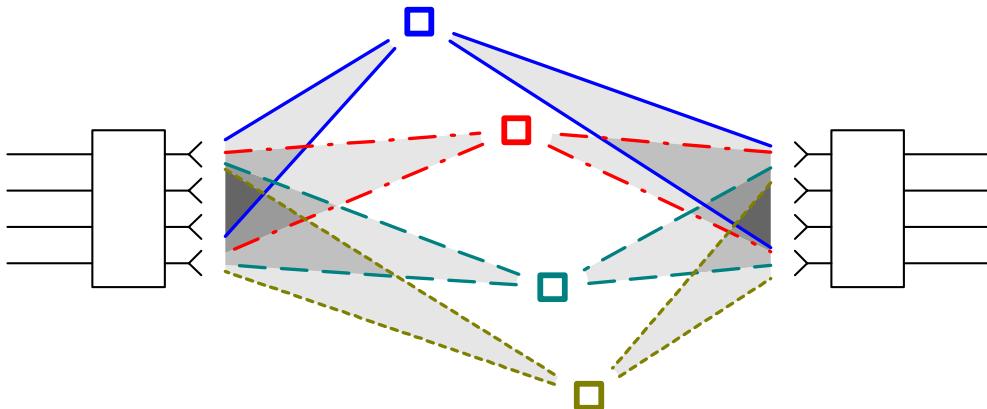


Figure 5-2: Creation of spatial sub-channels in MIMO systems by beam-steering

Apart from the capacity enhancement, the MIMO channel, in principle, also provides a diversity advantage. The elements of the channel matrix can potentially each provide a signal, and hence, provided they each fade independently; the available diversity order is $n_T \times n_R$. This is referred to as *full diversity*. To achieve full diversity in a practical system requires careful signal design.

Clearly the performance of MIMO systems, both in terms of capacity and diversity, depends strongly on the propagation environment. In some cases, where there is a strong line of sight and very little multi-path, MIMO gives no advantage because there is no spatial diversity available. It has been traditional to evaluate the performance of MIMO systems assuming uncorrelated Rayleigh fading between each pair of transmit/receive antennas, but this, in effect, assumes an infinitely rich multi-path environment. In practice, both capacity and diversity are limited by the number of multi-path components.

The discussion above shows that the number of sub-channels, and hence the capacity enhancement, is limited by the number of multi-paths; the diversity order is similarly limited. We shall consider below the effect of more realistic channel models. We shall also consider the case in which the channel is known, at least to some degree of accuracy, at the transmitter. This allows the transmitter to adapt to the channel, which is well suited to an SDR system. We shall demonstrate some of the potential advantages available in such systems.

As mentioned above, there are two main types of transmission scheme for MIMO systems, historically having their origins in independent work at AT&T and Lucent Bell Laboratories, respectively. The former, described as space-time codes, were developed out of existing diversity techniques and are designed to maximise diversity and, in particular, to achieve *transmit diversity* when multiple antennas are available only at the transmitter. The original paper on the subject [124] focuses on criteria to ensure full diversity in a space-time code. As mentioned above, two types of space-time codes were developed – STTC and STBC – but the latter have become the most popular because of their ease of decoding and ease of adaptation to various wireless standards.

UNCLASSIFIED

The second type of scheme, spatial multiplexing, was motivated by the capacity gains available. The techniques involved are closely related to multi-user detection, in which signal processing at the receiver is used to separate information streams transmitted over the same wireless channel. Significant diversity advantages are not usually obtained, although it is possible to design schemes to achieve nearly full diversity, borrowing the theoretical framework and the design criteria from space-time codes [125]. Again there are several forms, of which one, V-BLAST [122], has become most popular because of its ease of implementation, even though it does not achieve any significant diversity advantage, and it requires at least as many receive as transmit antennas.

5.2.1 MIMO Channel Models

As mentioned previously, MIMO systems were originally assessed assuming independent Rayleigh fading between each pair of transmit/receive antenna elements. However, it has been known for some time that this model is likely to yield optimistic estimates of performance. In effect, it assumes that there are a very large number of multi-paths, which of course is not the case in all environments. The discussion above has already shown that if there are a limited number of multi-paths both capacity and diversity gains will be limited.

More realistic models will account for the actual multi-path components, defining the direction of departure (DoD) from the transmitter, the direction of arrival (DoA) at the receiver, the attenuation and phase shift of the component and the time delay. (The latter can be neglected in the case of a narrow-band system, subject to flat fading.) Such models may be described using the framework of the *finite scatterers model* [126], which is shown in Figure 5-3.

The fundamental assumption here is that the radio propagation environment can be modelled in terms of a finite number of discrete propagation paths. Each path has a unique DoD, $\phi_{T,p}$, DoA, $\phi_{R,p}$, complex path gain, ξ_p , and delay, τ_p ($p = 1 \dots n_s$), as illustrated for path (a) in Figure 5-3. Note that the model is not restricted to 'single bounce' paths, as shown by path (b). 'Split' paths like (c) can be accommodated in the model by treating them as two paths which have the same DoD but different DoAs. The DoDs and DoAs are usually azimuth angles, but the model can readily be extended to include elevation.

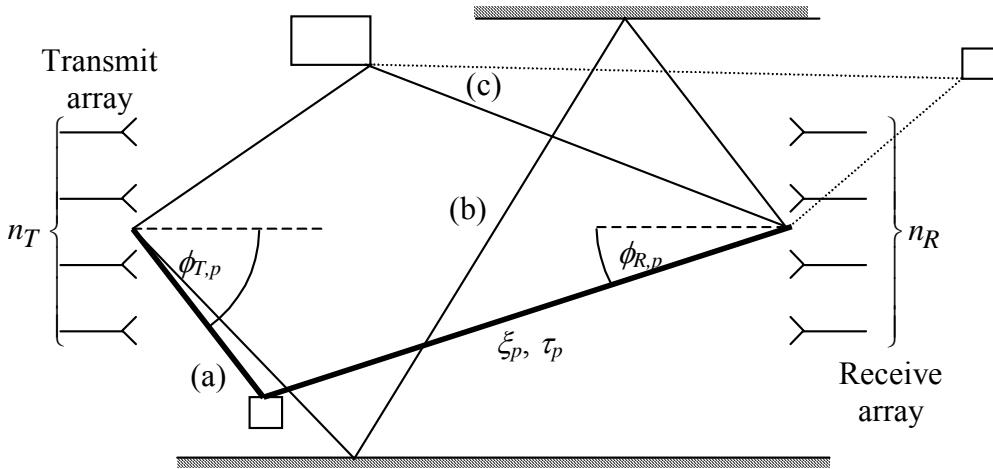


Figure 5-3: Finite scatters model, showing (a) DoD, DoA, path gain and delay of a typical path; (b) paths subject to multiple reflections; (c) 'split' paths with more than one DoA corresponding to a DoD

Here, \mathbf{H} is given by:

$$\mathbf{H} = \sum_{p=1}^{n_s} \xi_p \Psi_R(\phi_{R,p}) \Psi_T^T(\phi_{T,p}) = \Psi_R \Xi \Psi_T^T \quad \text{Equation 5-5}$$

where Ψ_R and Ψ_T are the *steering vectors* of the receive and transmit arrays, respectively; that is, the vector of signals on the elements when illuminated by a multi-path from the direction ϕ . In terms of the element positions, the steering vectors are:

$$\Psi = \left\{ \exp\left(2\pi j \frac{x_i \cos \phi + y_i \sin \phi}{\lambda}\right), i = 0 \dots n-1 \right\} \quad \text{Equation 5-6}$$

where \mathbf{x} and \mathbf{y} are vectors giving the x and y coordinates of the elements, ϕ is DoD/DoA, λ is the wavelength and n is the number of antenna elements. Note that this assumes ideal, uncoupled omni-directional elements. This can be extended to the three dimensional case:

$$\Psi = \left\{ \exp\left(2\pi j \frac{(x_i \cos \phi + y_i \sin \phi) \cos(\varphi) - z_i \sin(\varphi)}{\lambda}\right), i = 0 \dots n-1 \right\} \quad \text{Equation 5-7}$$

where additionally \mathbf{z} is a vector of z coordinates for the elements, and φ is the DoD/DoA in elevation. Except where there is reason to do otherwise, we will consider horizontal uniform linear antenna arrays (ULAs) in the sequel; in this case the steering vector becomes:

$$\Psi = \left\{ \exp\left(2\pi j \frac{il}{\lambda} \sin(\phi)\right), i = 0 \dots n-1 \right\} \quad \text{Equation 5-8}$$

where l is the spacing between antenna elements.

For given array geometries and given distributions of DoA, etc., this model can be used to obtain randomly-chosen instances of MIMO channel matrices. Commonly chosen distributions of DoA/DoD are uniform around 2π , or showing a Laplacian distribution centred on a particular direction:

$$P(\phi) = \frac{c}{\sigma_\phi \sqrt{2}} \exp\left(-\frac{|\phi - \phi_0| \sqrt{2}}{\sigma_\phi}\right) \quad \text{Equation 5-9}$$

where σ_ϕ is the RMS angular spread of the distribution, ϕ_0 is the centre of the distribution and c is a constant to maintain normalisation. We will assume that the path gains, ξ , have a Rayleigh distribution, and normally that their RMS amplitudes are all equal.

5.3 Literature Review

5.3.1 SDR Implementations of MIMO Systems

Individually, SDR and MIMO are very active research areas, and have been so for nearly a decade in both cases, generating an enormous amount of literature. It is perhaps surprising, therefore, that a literature search on the combination of the terms produces very few references indeed – a total of eight. Appendix A contains all the relevant abstracts obtained from a number of joint searches on related keywords in several bibliographic databases. An Internet search (using the Google search engine) of course produced a very large number of hits, but nearly all of these were simply documents in which both terms happened to occur, rather than documents dealing with the combination of the technologies. Appendix B also contains the first few pages of the Google results. A search for documents on the SDR Forum web site [24] using the keyword “MIMO” also gave no hits.

The references found are of two types. Three are general, one [127] giving a general vision of wireless communications which includes MIMO and SDR, one [128] describing a testbed for MIMO SDR systems and one [129] considering analogue-controllable RF devices for SDR-MIMO. The remainder address specific problems related to MIMO systems, giving solutions which can be implemented in SDR. References [130], [131] and [132] consider MIMO-OFDM systems. References [130] and [133] consider ‘smart antenna’ implementation and applications.

UNCLASSIFIED

This paucity of reported work suggests that MIMO has not previously been regarded as a particular issue for SDR, but simply as another set of functions which need to be implemented in software. In this report, however, we revisit this question from two points of view: firstly of the barrier it may present due to the additional hardware (including RF hardware) and software complexity involved, and secondly of the advantages SDR might bring to MIMO by enabling adaptivity. While adaptive MIMO is not new, this particular connection does not seem to have been made previously.

5.3.2 MIMO in Existing and Proposed Standards

The second area of ‘prior art’ we should consider is the use of MIMO in wireless standards, since, in the medium term at least, these are what future SDR systems will be called upon to implement. The wireless standards that are currently most important, including GSM, UMTS, Bluetooth wireless technology and IEEE 802.11, were all defined in their current form too early to incorporate MIMO to a very significant degree. However, given the effectiveness of MIMO in increasing capacity and diversity, and hence both power and bandwidth efficiency, there has been much work recently in the standardisation bodies to incorporate MIMO. Thus, while the release of the UMTS specification on which current implementations are based, Release 99, incorporated only two simple approaches to dual-antenna transmit diversity, more recent issues have included extensions to more than two antennas. In particular, work on the proposed increased data rate enhanced downlink of UMTS for multimedia applications, known as high-speed downlink packet access (HSDPA), includes several approaches for the use of MIMO to enhance capacity, not just transmit diversity to enhance robustness. UMTS specifications and working documents are available on the 3GPP web-site [134].

Similarly, the latest version of the IEEE 802.11 WLAN standard currently under development, 802.11n, is certain to include MIMO techniques to increase data rate (if only because this is the only feasible way of reaching its data rate target of 500 Mbit/s within the frequency allocations available). At the time of writing, there are four fully-developed proposals for the standard being considered by the standards committee, of which two, known as WWiSE and TGnSync, are front-runners. We will review these below. Again, all documents are available on the Internet [135].

5.3.2.1 UMTS

As mentioned previously, Release 99 of the UMTS specification included two forms of transmit diversity; an open loop form [136] called space time transmit diversity (STTD) and a closed loop form [137] called time switched transmit diversity (TSTD). Both are applied at the base station so that second order diversity can be provided on the downlink, given that many base stations (even for 2G) already employ dual diversity antennas in each cell sector. These can of course be used to provide conventional second order receive diversity on the uplink.

STTD is based on the Alamouti dual antenna transmit diversity scheme [138]. The principle here is to map two modulated data symbols to two transmit antennas such that the transmitted signal can be written:

$$\mathbf{S} = \begin{pmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{pmatrix} \quad \text{Equation 5-10}$$

in which the columns represent signals transmitted on the same antenna, while the rows are transmitted in the same symbol period. In STTD, however, this is inverted such that the rows are transmitted on the same antenna, the columns at the same time. The advantage of this scheme is that it can provide optimum performance using a very simple linear detector, and can achieve full diversity. The primary application is to provide second order diversity on the downlink even when only one receive antenna is available on the terminal. Note that since the transmitter does not depend on knowledge of the channel to achieve this diversity, it remains effective even if the channel is quite rapidly time variant. In fact, it remains effective until the Doppler frequency becomes an appreciable fraction of the symbol rate.

TSTD is a closed loop scheme which essentially adjusts amplitude and phase of the signals on the two transmit antennas at the base station in such a way as to maximise the signal at the user terminal. Figure 5-4 shows the signal processing at the transmitter for a particular user. After spreading and scrambling by the appropriate channelisation and scrambling sequences, the data are split and each branch is multiplied by a different weight, which, in general, is complex. A different pilot sequence is added to each branch, which is then transmitted on one antenna (this is the conventional CPICH channel). The weights are determined by feedback from the user terminal, where the channel from each antenna is estimated using the pilot from that antenna, and the optimum transmitter weights chosen. The selection is then signalled back to the transmitter by an uplink control channel. There are two versions, depending on the possible set of weights. In the simplest version, the weights correspond to a phase shift of zero or π , and the terminal simply chooses the phase which will result in the largest signal at its antenna. In the second version, a much wider range of weights is available.

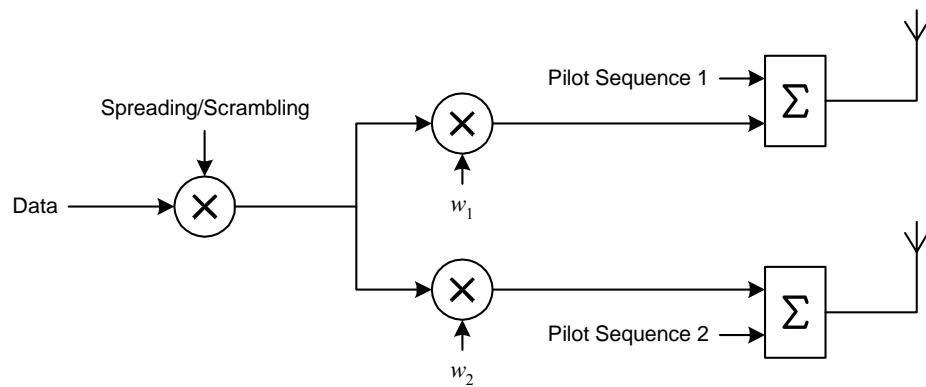


Figure 5-4: Transmission for TSTD technique

UNCLASSIFIED

In this scheme there will be a delay of at least one slot period, possibly several, between the channel estimation and the choice of weight at the transmitter. Hence the technique is suited only to cases where the channel changes relatively slowly; the maximum Doppler must be much less than the slot rate. Note however that the scheme is adaptive: not only does it provide a diversity advantage, which even in the simpler version amounts to selection diversity, but it also provides a very simple form of beamforming which increases the gain in the direction of the wanted user and reduces the interference to other users.

In a more recent release of the specification [139] a number of new concepts have been introduced to extend transmit diversity to more than two transmit antennas. For example, the closed loop TSTD technique has been extended by the use of an *eigenbeamforming* concept (see Section 5.2, and further discussion below), since otherwise this would tend to require much more feedback. The eigenvectors \mathbf{V} and \mathbf{U}

of the long-term average channel autocorrelation matrix, \mathbf{HH}^H , are calculated (this need only be done relatively infrequently) and the strongest are selected. Weights are applied at the transmitter corresponding to these modes. To compensate for fast fading, the strongest eigenmode is selected at each slot period and the appropriate set of weights selected. There are also open loop extensions to the STTD technique, even though the Alamouti scheme cannot be used directly for more than two antennas: Reference [124] shows that either the orthogonality that gives rise to the simple decoder, or the throughput rate, must be sacrificed. However [139] includes two pure open loop concepts relying on randomisation which achieve the required effect. In one, a second pair of antennas is fed via a time-varying pair of phase rotations; in the other a 4×4 full rate space-time block code is used along with symbol level scrambling. There is also a closed loop version which also uses a second pair of antennas fed via phase rotations, in this case selected by feedback from the user terminal in essentially the same way as the two antenna closed loop system. All these, however, appear to have the status of proposals for future releases rather than the current specification.

5.3.2.2 HSDPA

The HSDPA extension to UMTS currently under discussion in 3GPP includes several proposals for the use of MIMO techniques. Those being most actively considered are described in a report on the relevant work item [140]. This is evidently a ‘work in progress’, since many of the subsections are empty in the latest version available. Since the purpose of HSDPA is to increase the downlink data rate for multimedia applications, the focus is on true MIMO techniques, with multiple antennas at transmitter and receiver, and hence on spatial multiplexing rather than space time coding, or other transmit diversity techniques. There are a total of eight proposals: it is clearly impossible to describe them all in detail here. However, they can be divided into several groups.

Proposals 1, 6 and 7: per-antenna rate control (PARC); TPRC for CD-SIC MIMO; selective per-antenna rate control (S-PARC) can be described as rate-controlled spatial multiplexing scheme: the archetype is probably PARC, shown in Figure 5-5 (diagrams in this section are taken from [140]).

UNCLASSIFIED

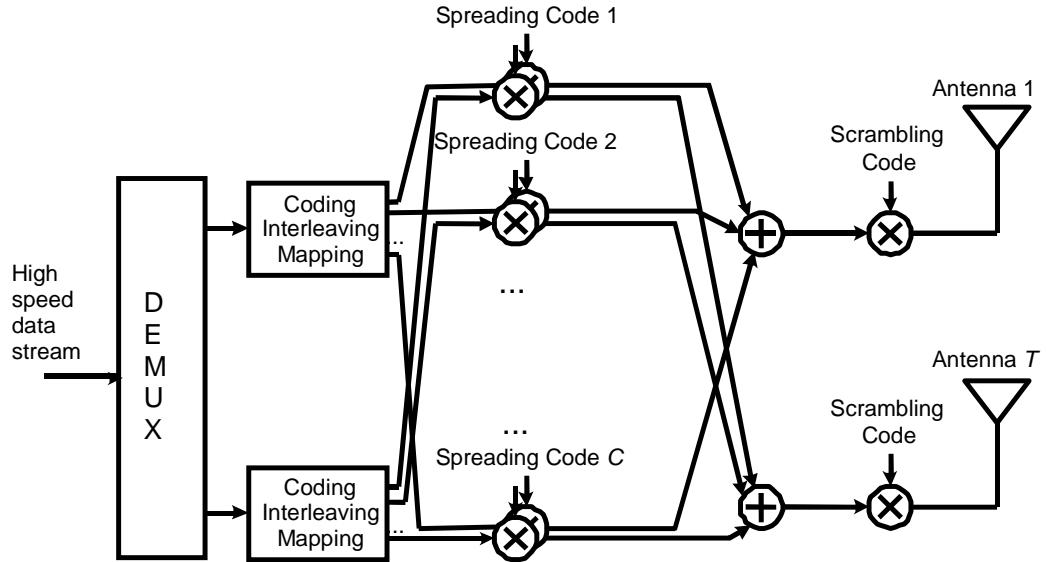


Figure 5-5: PARC scheme (diagram taken from [140])

The data are demultiplexed into a stream per antenna, possibly of different data rates, which are separately coded on each antenna (and may also be power weighted). The stream may then be further demultiplexed into several streams, corresponding to the use of multiple orthogonal variable spreading factor (OVSF) spreading sequences per user to provide a higher data rate, then spread per code and scrambled per antenna. There are different options on the use of spreading and scrambling codes to optimise capacity. Additionally, the S-PARC scheme allows the use of a subset of the available antennas to be selected (which only takes rate control to its logical conclusion).

Proposals 2 and 3: rate-control multi-paths diversity (RC MPD); double space time transmit diversity with subgroup rate control (DSTTD-SGRC) are hybrids of spatial multiplexing and STTD: the transmit antennas are divided into multiple groups of two, to which the STTD scheme described in the previous section is applied. Then the data are spatially multiplexed and separately coded across these groups in a similar way to the PARC proposal over individual antennas. Figure 5-6 describes the latter scheme, which is similar to the former.

UNCLASSIFIED

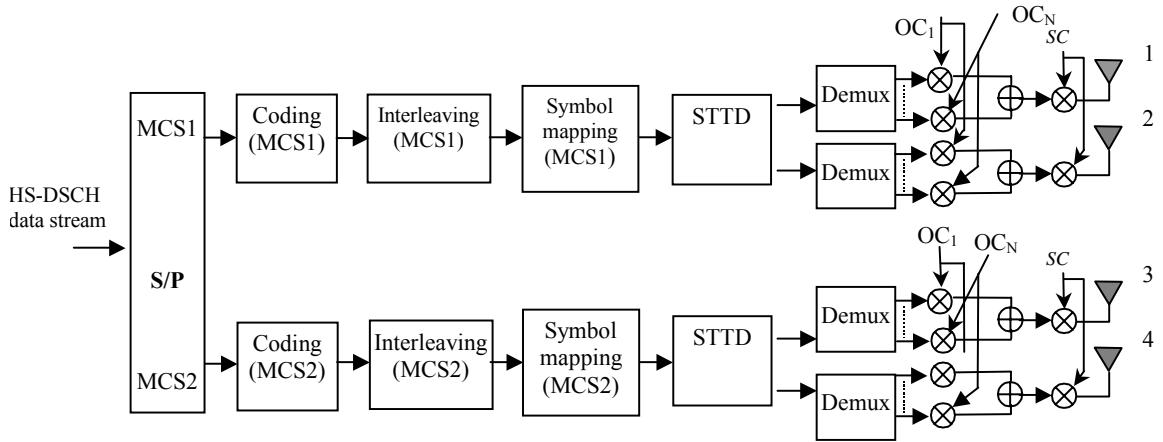


Figure 5-6: Illustration of double space time transmit diversity with subgroup rate control (DSTTD-SGRC), (diagram taken from [140])

Similarly proposals 4 and 8: single stream closed loop MIMO with four transmit and L receive antennas; double transmit antenna array (D-TxAA) are hybrids of spatial multiplexing and the TSTD closed loop transmit diversity technique of the previous section. Figure 5-7 illustrates the scheme, showing the complex weights on each antenna which are adjusted according to feedback from the mobile terminal.

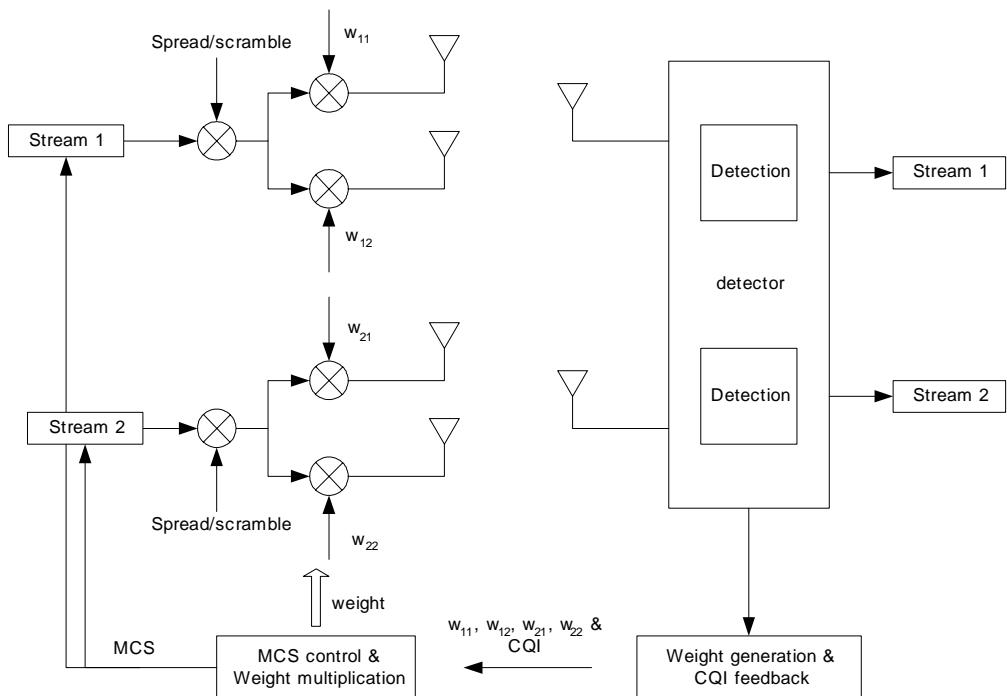


Figure 5-7: Illustration of double transmit antenna array scheme (diagram taken from [140])

Finally proposal 5: per-user unitary rate control (PU^2RC) uses the singular value decomposition to perform fully-adaptive MIMO transmission. After demultiplexing into multiple streams, again possibly of different rates, encoding, spreading, scrambling, etc., the data is transformed by a unitary matrix, effectively the matrix \mathbf{U} in Equation 5-4. Thus the data is transmitted on the channel eigenmodes, which can be shown to provide the optimum capacity if knowledge of the channel is available at the transmitter. Hence the scheme requires feedback from the transmitter of the exact channel state. The scheme is illustrated in Figure 5-8.

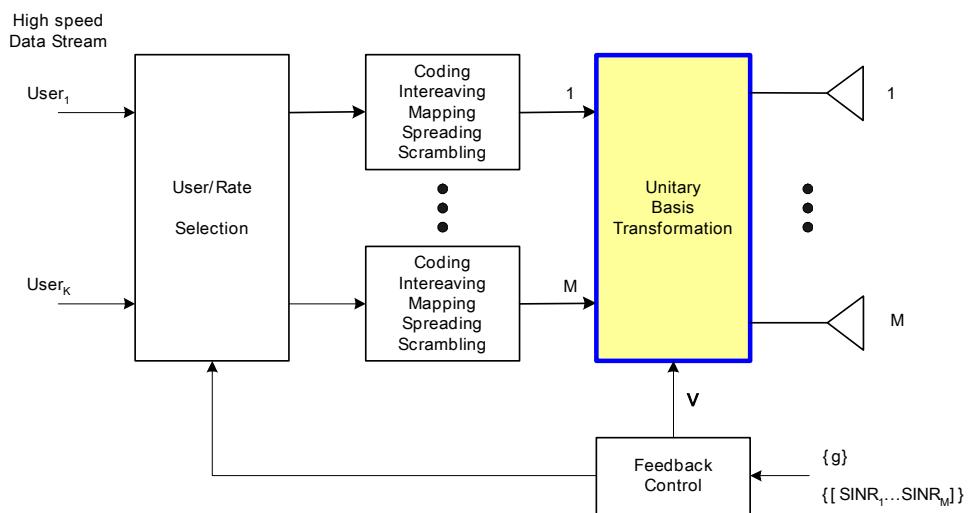


Figure 5-8: Illustration of per-user unitary rate control (PU^2RC), (diagram taken from [140])

5.3.2.3 IEEE 802.11n

As mentioned above, standardisation of the high-speed WLAN standard, IEEE 802.11n is still in progress, but it is certain to include MIMO since that might be described as the main reason for a new standard being developed. Given the bandwidth available, and that practical considerations restrict the size of the constellation, MIMO was the only way to achieve a doubling of the data rate over the 54 Mbit/s of IEEE 802.11a/g.

At the time of writing, two proposals for 802.11n can be identified as ‘front runners’, namely, ‘TGn Sync’ and ‘WWiSE’. These have much in common, most fundamentally that MIMO is central to their approach and their physical layer is based on OFDM. Figure 5-9 and Figure 5-10 show the block diagrams of the transmitters of the two proposals, respectively, taken from the presentations of the full proposals given to the standardisation committee [141], [142]. Apart from the format in which they are presented, these are nearly identical in structure: they employ spatial multiplexing between two transmit antennas. In fact the main difference is that TGn Sync proposes that a mandatory 40 MHz transmission bandwidth should be possible for all equipment, as well as 20 MHz, even though it could not be used in all regulatory regimes. WWiSE would make only 20 MHz mandatory. 40 MHz would of course allow a higher data rate with fewer transmit antennas, and therefore smaller terminals.

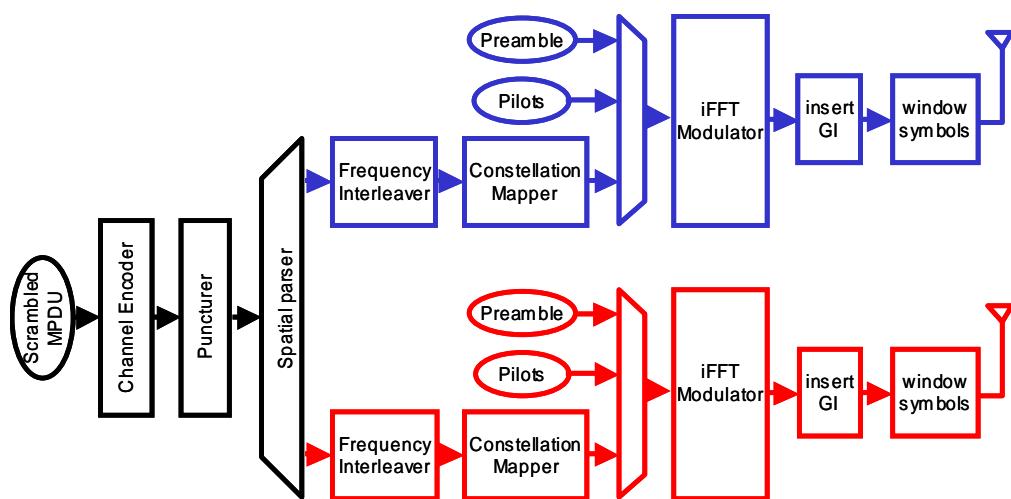


Figure 5-9: Transmitter block diagram of TGn Sync proposal, taken from [141]

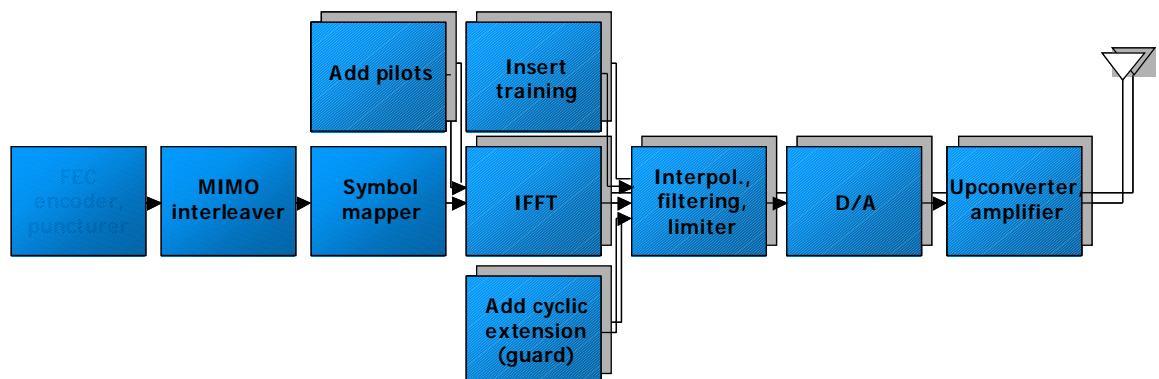


Figure 5-10: Transmitter block diagram of WWiSE proposal, taken from [142]

UNCLASSIFIED

As mentioned above, for both proposals the basic, mandatory structure supports spatial multiplexing over two transmit antennas, and hence requires two receive antennas. The data are encoded, interleaved and de-multiplexed between the two antennas in much the same way in both. Spatial multiplexing over more than two antennas is also possible when the same number of antennas is available at transmitter and receiver. The use of MIMO where the number of transmit and receive antennas differ is in the optional extensions which apply in asymmetric situations, when more transmit than receive antennas are available. TGn Sync proposes *orthogonal spatial spreading* in which the two streams are each applied to all transmit antennas after weighting by a different set of orthogonal antenna weights. This is an open loop scheme, requiring no feedback about the channel state. It also allows closed loop transmit beamforming, if such feedback information is available. WWiSE, on the other hand, uses a hybrid of space time block codes and spatial multiplexing in this situation, and has no option for a closed loop technique.

5.4 RF Architectures for MIMO Implementation in SDR

If MIMO is deemed essential for SDR and it turns out that MIMO is either incompatible with SDR, or it is excessively expensive, then this could be a barrier to future wireless systems. There are two aspects to the implementation: RF issues, including antennas, and baseband signal processing. In this subsection we consider RF subsystems.

The main issue here for MIMO is simply that MIMO inherently requires multiple antennas, which suggests multiple RF chains and ADCs/DACs. In one sense this is irrelevant to the SDR issue since any MIMO system would inherently require these. However, there may be concerns if the terminal is required to reconfigure between different standards or modes which involve different numbers of antennas. It would then be desirable to avoid the need to supply RF chains which are only used in one mode. Antenna arrays are a particular concern if different standards operate in different frequency bands. Moreover, if antenna elements can be used in only one frequency band, then MIMO would require a total number of elements (and RF chains) given by the number of bands of operation multiplied by the number of elements in the array, which would rapidly become infeasible.

For this reason in this section we consider the implementation of RF subsystems for MIMO systems, including multi-band antenna arrays, and ways of reducing the requirement for RF hardware, especially the possibility of sharing RF chains.

5.4.1 Antennas

There are two questions concerning antenna arrays for MIMO systems operating in different frequency bands. The first applies to the antenna elements in isolation: to what extent can they operate efficiently in two or more different bands? The second applies to the array: since the optimum spacing of a MIMO array often depends on wavelength, can an array provide good MIMO performance at several different wavelengths?

UNCLASSIFIED

These questions have been extensively analysed by the European FLOWS IST project, in particular in the FLOWS report D16 [143], some of whose results are drawn upon here. FLOWS has considered the construction of arrays of multi-band antennas, specifically for the three bands covering GSM 1800, UMTS (around 2 GHz) and IEEE 802.11a (around 5 GHz). Two approaches to such multi-band antennas have been considered: triple-band, covering the three bands separately (or possibly in two bands), and UWB, covering the entire bandwidth and including all three bands. There is an inherent disadvantage in all these approaches in that the elements are required to be larger than a single band antenna would be. This is especially the case for the UWB antennas. Figure 5-11 and Figure 5-12 show examples of these two types of antenna, along with their frequency responses. This shows that multi-band antennas of this type can be implemented at a size which makes them feasible.

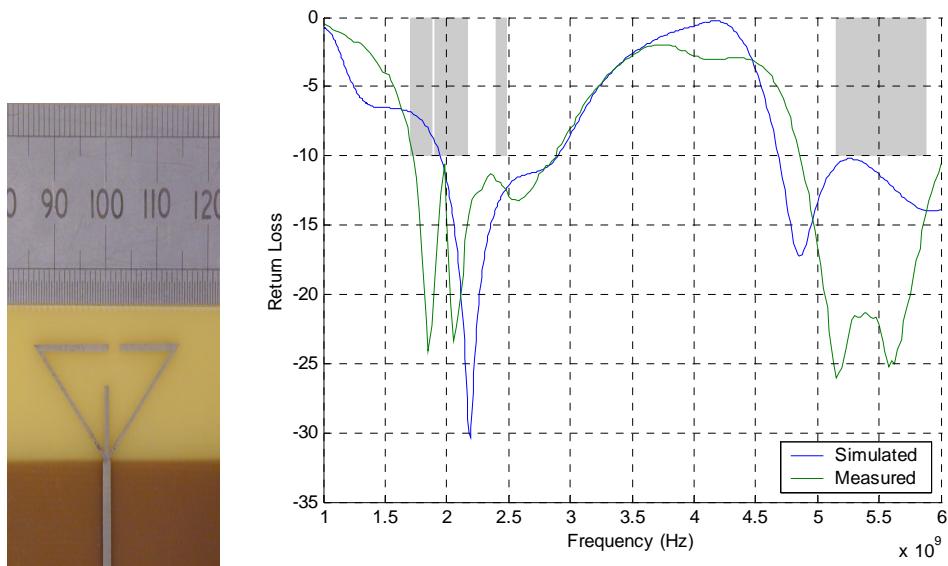


Figure 5-11: Three branched monopole triple-band antenna (left) with simulated and measured return loss (right), taken from [143]

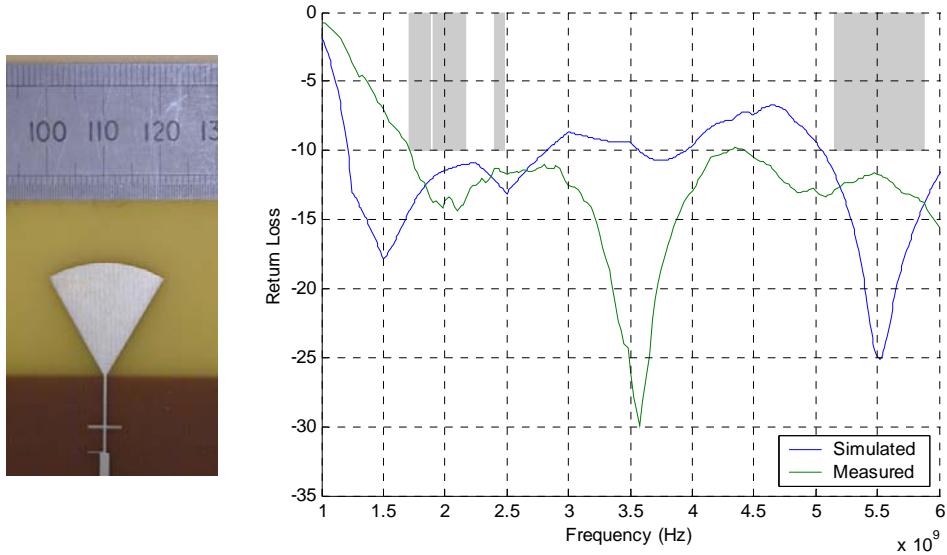


Figure 5-12: UWB 'half bowtie' antenna (left) with simulated and measured return loss (right), taken from [143]

The second question concerns the spacing of the elements in order to provide maximum benefit. A guide to this may be obtained using the finite scatterers channel model, as described in Section 5.2.1, to estimate the distribution of the MIMO capacity given by Equation 5-2. Since the capacity depends on the correlation of fading between the antenna elements, the correlation function of signals across the array also provides a useful indication. It may be shown [144] that the correlation between two antenna elements spaced by l_{ik} is given by:

$$\overline{R_{R,ik}} = n_T \int_{-\pi}^{\pi} \zeta(\phi_R) \exp j(2\pi \sin(\phi_R) l_{ik} / \lambda) d\phi_R \quad \text{Equation 5-11}$$

where $\zeta(\phi_R)$ denotes the density function of path arrivals, defined as:

$$\begin{aligned} \zeta(\phi) &= \lim_{\delta\phi \rightarrow 0} \left\{ \frac{1}{\delta\phi} E \left[\sum_{p, \phi \leq \phi_{R,p} < \phi + \delta\phi} \overline{|\xi_p|^2} \right] \right\} = \lim_{\delta\phi \rightarrow 0} \left\{ \frac{1}{\delta\phi} \sum_{p=1}^{n_S} \overline{|\xi_p|^2} P(\phi \leq \phi_{R,p} < \phi + \delta\phi) \right\} \\ &= \lim_{\delta\phi \rightarrow 0} \left\{ \frac{1}{\delta\phi} \sum_{p=1}^{n_S} \overline{|\xi_p|^2} p(\phi_{R,p}) \delta\phi \right\} = \sum_{p=1}^{n_S} \overline{|\xi_p|^2} p(\phi_{R,p}) \end{aligned} \quad \text{Equation 5-12}$$

UNCLASSIFIED

Equation 5-11 can be regarded as a version of the Fourier transform of the arrival density function from the angular domain to the aperture domain, noting that the term within the exponential is $\sin(\phi_R)$ rather than ϕ_R . This shows that the correlation depends on the distribution of the AoAs. If this distribution is uniform in all directions, then elements will become uncorrelated, on average, at relatively small spacing, while if the AoAs are restricted to a narrower range, then the spacing must be large for low correlation. Note that if the distribution is uniform, the correlation as a function of spacing is given by a Bessel function:

$$\overline{R_{R,ik}} = J_0(2\pi l_{ik}) \quad \text{Equation 5-13}$$

where J_0 denotes the zero order Bessel function of the first kind. Figure 5-13 shows the resulting correlation function.

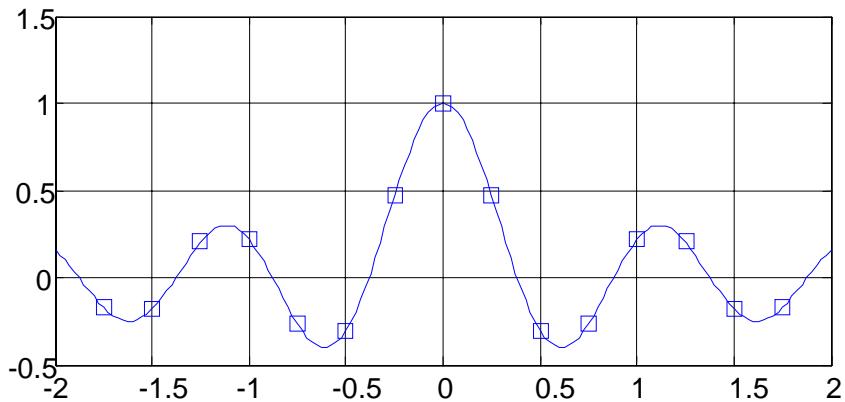


Figure 5-13: Correlation function of fading at antenna elements as a function of spacing in wavelengths, taken from [143]

Figure 5-14 shows the resulting capacity distribution for a system with four receive elements as a function of element spacing. It shows that capacity increases with spacing up to approximately $0.4\times\lambda$, but for larger spacing remains approximately constant. This suggests that a multi-band MIMO array should be designed such that the spacing is at least $0.4\times\lambda$ at the lowest frequency of operation.

UNCLASSIFIED

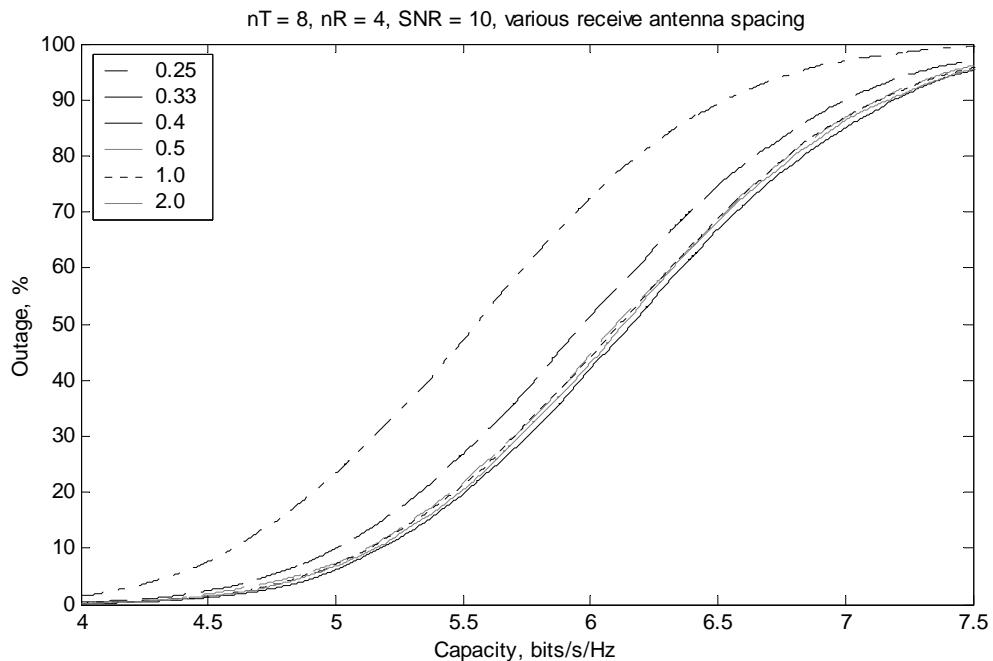


Figure 5-14: Capacity distribution for an 8 × 4 MIMO system, 64 scatters, with various receive antenna spacings (in wavelengths), taken from [143]

(Note that Figure 5-13 and Figure 5-14 contradict the usual wisdom that an ideal array is uncorrelated for a spacing of $\lambda/2$: in fact the minimum correlation is reached at 0.38λ , which is the first null of the Bessel function. However even this does not provide a null in the correlation of any array of more than 2 elements, since the subsequent nulls are not multiples of this, so it is impossible to design an uniform linear array with zero correlation between all elements. Nevertheless the effects of this on the capacity are small).

However this design rule would be likely to result in large arrays, whereas most terminals have limited size. Figure 5-15 shows the capacity distribution for the case where the total aperture of the array is limited by the size of the antenna to two wavelengths. Again, ideal uncoupled antenna elements are assumed. Because this simple model of an antenna array would otherwise result in a gain due simply to the increased power collected by a larger number of elements, the channel has been normalised to remove this array gain. Under these conditions we observe that for more than six antennas, corresponding to a spacing of 0.4λ , there is no additional capacity increase beyond that due to array gain.

UNCLASSIFIED

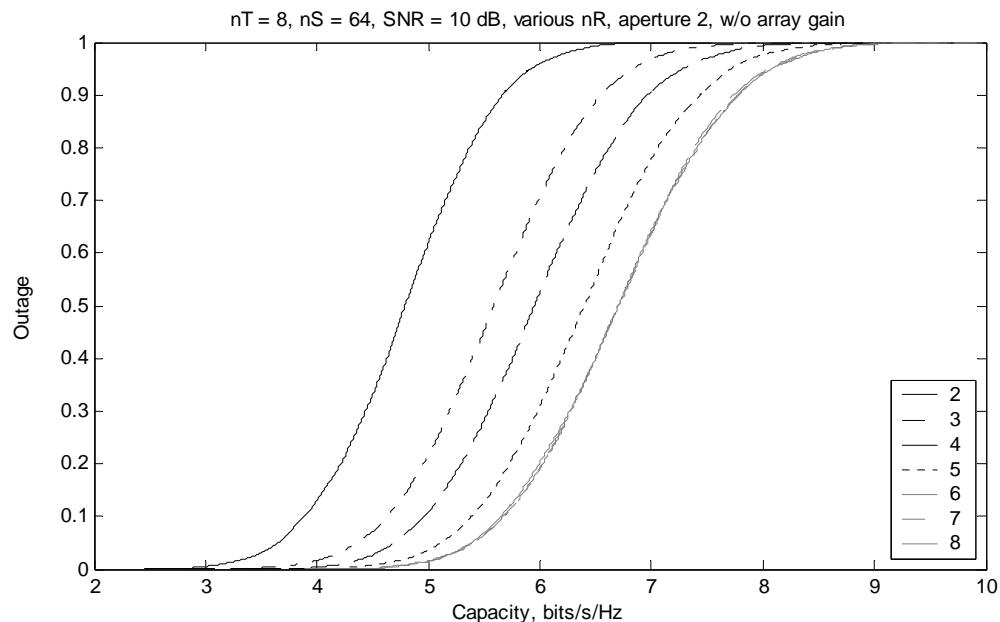


Figure 5-15: Capacity distribution for 8 transmit elements and a receive array with total aperture of two wavelengths with various number of elements, normalised to eliminate array gain

In view of this, one solution likely to be optimal where a terminal is required to handle systems in different frequency bands in a limited size, is to use a combination of multi-band and single band elements such that the minimum spacing between elements operating on a given band is $0.4 \times \lambda$ for that band. In this way the maximum benefit available is provided while minimising the hardware requirements, both in terms of antennas and RF hardware. Figure 5-16 illustrates this for an array handling signals at 900 MHz, 1.8 GHz and 5 GHz, limited to a terminal size of 7 cm. The resulting array supports one element at 900 MHz, two at 1.8 GHz and four at 5 GHz, so that the benefit of a MIMO terminal is available at the bands at which it provides a worthwhile benefit.

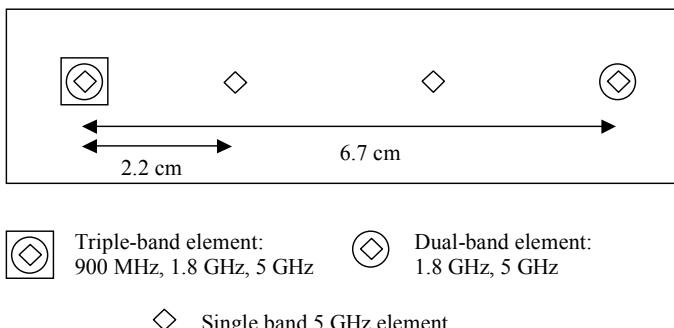


Figure 5-16: 7cm terminal with multi-band array, supporting one element at 900MHz, two at 1.8 GHz, and four at 5 GHz

UNCLASSIFIED

5.4.2

RF Hardware: Separate Implementation

In addition to antenna elements, a MIMO system requires in principle a separate RF chain per antenna element, including LNA, filters, mixers, IF amplifiers and ADC for reception and DAC and HPA for transmission. (It may be possible to share LOs between chains). In addition, where operation is required over several frequency bands, some separate items may be required for each band (for example RF filters). Clearly the most straightforward option is to provide a separate RF chain for each element. Techniques for reconfiguring these for different frequency bands are considered elsewhere in the report. Clearly, however, this could result in significant extra RF hardware requirements and costs for a MIMO system, and would be a barrier to MIMO SDR adoption. The approach described above in respect of antennas will minimise these extra costs, but they remain significant. For this reason we briefly consider in the following subsections techniques to reduce this complexity.

5.4.3

Analogue RF Implementation

An approach that would at least minimise the number of ADC/DACs required is to implement some of the MIMO processing in analogue RF hardware. Most of the operations in a MIMO receiver involve linear operations: multiplication by a set of complex weights followed by summation. These can in principle be implemented in analogue hardware, using digitally-controlled phase shifters and attenuators. It is again beyond the scope of this part of the report to consider these in great detail, but note that MIMO algorithms generally require very flexible and accurate operations to be carried out.

A configuration that might be considered for this purpose is the Butler matrix [145]. This is a network of mixers and phase shifters that effectively performs beamforming into a set of fixed orthogonal beams, corresponding to a Fourier transform of the signals on the elements. (This effectively transforms between the aperture domain and the angular domain.) For an n_R -element uniform linear receive array, the corresponding Butler matrix has n_R outputs, corresponding to n_R beams approximately evenly distributed in direction. This would have clear benefits for beamforming approaches if the required beams were close to those generated by the matrix, since the required output or outputs of the matrix could be selected, and a significantly reduced number of RF chains and ADCs could be used. (The same argument can be used regarding transmission.) However, it is likely that an LNA (at the receiver) and an HPA (at the transmitter) would be needed per antenna in any case. More significantly, MIMO techniques in general do not fit into this fixed beamforming scheme, and to extract all the received information (or to provide full flexibility on transmission) would require all outputs (or inputs) of the Butler matrix to interface with the baseband processing, which would provide no saving of RF hardware.

More generally, not all MIMO processing fits into the category of linear processing as described above. Notably most forms of spatial multiplexing, including V-BLAST, do not, since the fundamentally nonlinear process of interference cancellation is used. Moreover in the ordered successive interference cancellation (OSIC) algorithm of V-BLAST, the order of cancellation has to be changed ‘on the fly’, and this is unlikely to be feasible in an analogue implementation. STTCs use the Viterbi algorithm based on the received signals from all antennas and is thus not amenable of an analogue implementation. It is conceivable that STBC might be possible, since the main element of decoding is multiplication by a matrix, but this matrix is determined by the channel, and would require very accurate implementation. These considerations suggest that analogue implementations of MIMO algorithms in RF hardware would, in general, not provide a useful solution for MIMO SDR systems.

5.4.4 Multiplexed RF Chains

Recently the FLOWS project has considered RF hardware implementations of MIMO terminals [146], and has proposed an approach based on multiplexing of the RF chains. The output signals of the antenna elements are multiplexed together, and a single common RF chain and ADC is used for them all (and, once again, equivalently for the transmit direction). In principle, time-division, frequency-division or code-division multiplexing can be used, although the FLOWS proposal, now the subject of a patent application [147], uses code-division multiplexing, as shown in Figure 5-17. The binary phase-shift keying (BPSK) modulator modulates the local oscillator with four different orthogonal binary sequences, with a symbol rate four times the received symbol rate. The signals are then combined over a single IF chain and a pair of ADCs (for in phase and quadrature components), and the streams are separated in the baseband processing by correlating the signal with the four codes. An LNA and a mixer are still required per antenna (noting that the mixer shown replaces the RF front-end mixer that would be required for separate implementation), but the IF and the ADCs are combined. The disadvantage is that their bandwidth must be multiplied by the number of antennas.

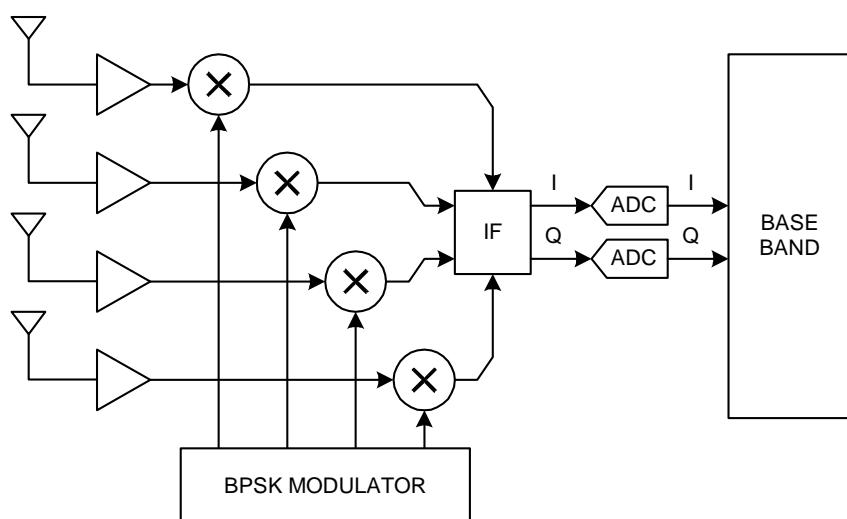


Figure 5-17: Multiplexed MIMO RF chains, taken from [146]

UNCLASSIFIED

This architecture is quite well suited to an SDR implementation, since not only does it save RF hardware, but it is highly flexible as to the frequency band, the signal bandwidth and the number of antennas used.

5.5 Outline of MIMO SDR Implementation

We consider next the implementation of the baseband signal processing and, in particular, the additional complexity resulting from the use of MIMO techniques. Reviewing the various techniques proposed for current or forthcoming standards in Section 5.3, we can identify three basic types of MIMO technique used, either separately or in combination, in these systems. These are STBC, spatial multiplexing and closed loop techniques, which include schemes like TSTD as well as more explicit beamforming techniques. (All such schemes involve the application of different weights at the transmitter, according to feedback from the receiver.)

Since in most cases the bulk of the computation is required at the receiver, we will concentrate on the implementation of this. For most of the schemes being considered in the standards, block diagrams have been included above, which we will not repeat here. Note that all these schemes rely on knowledge of the channel, if not at the transmitter then certainly at the receiver. This implies that channel estimation is required at the receiver: we will also consider the complexity of this for MIMO systems, over and above what would in any case be required for a single-input, single-output (SISO) system.

5.5.1 STBC

We will illustrate the implementation of STBC with reference to the Alamouti scheme [138] (STTD in UMTS), since it is the simplest. Other possible schemes are described in [124], as well as in appropriate standards documents: the principles of decoding are basically the same.

The Alamouti scheme for two transmit and one receive antenna can conveniently be represented by the matrix equation:

$$\mathbf{r} = \mathbf{Hx} + \mathbf{n} \quad \text{Equation 5-14}$$

where $\mathbf{H} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}$, $\mathbf{r} = \begin{bmatrix} r_1 \\ r_2^* \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and \mathbf{n} is the vector of received noise.

A maximum likelihood (ML) detector can be formed by multiplying the received signal by the transpose conjugate of \mathbf{H} :

$$\hat{\mathbf{x}} = \mathbf{H}^H \mathbf{r} = \mathbf{H}^H \mathbf{Hx} + \mathbf{H}^H \mathbf{n} = \left(|h_1|^2 + |h_2|^2 \right) \mathbf{x} + \sqrt{|h_1|^2 + |h_2|^2} \mathbf{n}' \quad \text{Equation 5-15}$$

UNCLASSIFIED

where \mathbf{n}' denotes the noise vector as transformed by a unitary transformation, which does not affect its mean power. The outputs of this multiplication, which are estimates of the original transmitted data symbols, can then be demodulated in the same way as in a SISO system.

Hence, the additional computation required for the basic detection process is this matrix multiplication, i.e., four complex multiplications and two complex additions for two data symbols. To generalise this, an STBC of length m , transmitting k data symbols over n_T transmit elements requires $n_T \times m$ complex multiplications and $n_T \times (m-1)$ complex additions per k data symbols. Each data symbol can of course carry several information bits. Let us assume that QPSK is used: then for the Alamouti scheme only one complex multiplication and half a complex addition is required per bit, which is negligible.

In addition, as mentioned above, the channel (i.e., the values of h_1 to h_{nT}) must be estimated. This is more difficult than the equivalent for a SISO channel, since there are nT variables to estimate in place of one. It is usually done using pilot or ‘training’ symbols embedded in the transmitted signal. In the case of MIMO systems, these should ideally take the form of orthogonal sequences transmitted on each antenna element, since this allows optimum joint estimation of the channels from each transmit element. The length, N , of these must be at least n_T . In general, the estimation involves a matrix multiplication of the corresponding ($n_R \times N$) received signals by the ($n_T \times N$) set of pilot sequences: a total of $n_R \times nT \times N$ complex multiply-accumulate operations: $n_R \times n_T^2$ if the shortest pilot sequences are used, and n_T^2 if there is only one receive antenna (often the case for STBC). Since this is carried out once per frame, which is likely to include more than 100 data bits, in terms of operations per bit it is negligible. It is also readily carried out in pure software, and thus can be readily implemented in SDR.

5.5.2

Spatial Multiplexing

In spatial multiplexing schemes, essentially independent data streams are transmitted from each antenna element. These are separated at the receiver using techniques related to multi-user detection (MUD). A wide range of such techniques are available, from linear techniques based on zero-forcing (ZF) and minimum mean square error (MMSE) transformations, through interference cancellation to maximum likelihood sequence estimation (MLSE). The complexity of the latter increases exponentially with the number of transmit elements, and may therefore be infeasible for large transmit arrays. A combination of linear and interference cancellation approaches known as OSIC is proposed for the V-BLAST transmission technique [122], but the performance is in fact comparable with MMSE only, and therefore we will use this as a baseline for complexity computation here.

MMSE detection again involves multiplication of the received signal vector by a transformation matrix:

$$\hat{\mathbf{s}} = (\mathbf{R}_{HH} + \sigma^2 \mathbf{I})^{-1} \mathbf{H}^H \mathbf{r} \quad \text{Equation 5-16}$$

UNCLASSIFIED

where $\mathbf{R}_{HH} = \mathbf{H}^H \mathbf{H}$ and σ^2 is the variance of the noise. Note that this reduces to ZF (multiplication by the inverse of the channel matrix) when the signal to noise ratio is very high (i.e., when σ^2 is negligible), and to matched filtering (multiplication by \mathbf{H}^H) when it is very low (i.e., when σ^2 is very large). The overall transformation matrix is $(n_T \times n_R)$: the multiplication requires $n_T \times n_R$ complex multiplications and $n_T \times (n_R - 1)$ additions for n_T data symbols. In this case usually $n_R = n_T$, so this is n_T multiplications and $(n_T - 1)$ additions per symbol. If we assume a 4×4 MIMO system using QPSK, this is two multiplications and one and a half additions per bit: still negligible. Note that, however, the OSIC algorithm would be significantly more complex, since it requires in principle for the transformation matrix itself to be recalculated several times during the processing of one symbol.

This technique also requires the channel to be estimated. This will usually be more complex again than for STBC, since the channel now involves $n_T \times n_R$ variables. From the argument above the complexity will be $n_R \times n_T^2$ complex operations if the shortest pilot sequences are used: still negligible. In addition, it requires the matrix inverse in Equation 5-16 to be computed. Given that \mathbf{R}_{HH} is Hermitian, this can be done using Cholesky factorisation, with complexity approximately $n_T^3 / 6$ complex multiply-accumulate operations. For $n_T = 4$, this is 11 complex operations. This also need only be performed once per frame (since it depends only on \mathbf{H} and not on the data), and hence is also negligible.

5.5.3 Wideband Systems

The description above applies to narrowband systems, operating on frequency non-selective fading channels. Most practical systems are wideband, in that the transmission bandwidth is significantly greater than the coherence bandwidth of the channel, and hence the channel is frequency selective over the signal bandwidth, and causes signal dispersion. While this could be handled by a conventional single carrier system using equalisation, in practice most wireless systems use either CDMA or OFDM for such channels. Hence we need to consider the implementation of MIMO techniques in such systems.

In this case the channel estimation complexity is also increased, since in general each element of the channel matrix \mathbf{H} must be modelled as a tapped delay line of length equal to the product of the maximum delay and the bandwidth (say L), and all tap weights must be estimated. This involves in addition a multiplication by an $(L \times L)$ matrix, so the complexity becomes $n_R \times n_T \times L \times (N + L)$ complex operations. In this case, N will need to be at least equal to L , and L may be 16 in a WLAN system with a 20 MHz bandwidth. Then for, say, a 4×4 system, the complexity becomes $4 \times 4 \times (16 + 16) = 8192$ complex operations per frame, which is no longer negligible, although probably still acceptable in the context of a broadband system with slowly varying channels.

5.5.3.1 CDMA

CDMA systems deliberately spread the signal bandwidth so that it is significantly greater than the coherence bandwidth. This allows them to exploit frequency diversity on fading channels, by means of a RAKE receiver [148], shown in Figure 5-18. It consists of a series of L correlators, sometimes referred to as ‘fingers’ or ‘taps’, each fed by the same spreading code via a different delay, τ_i , $i = 1 \dots L$. The outputs are then weighted by complex weights w_i and combined. The principle is that the delays correspond to the delays of multi-path components, so that each correlator extracts one component, and these are combined by maximal ratio combining. In this way the receiver exploits the multi-path to increase diversity.

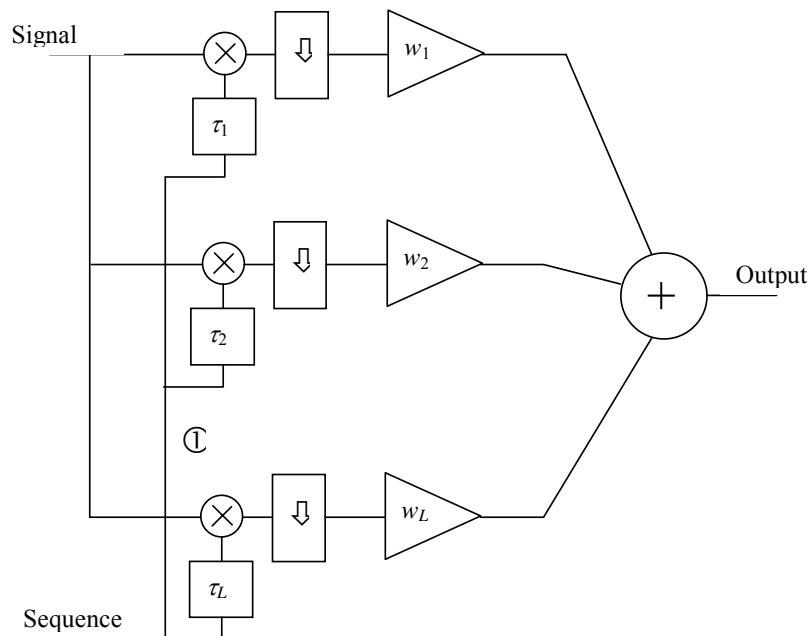


Figure 5-18: RAKE receiver

In a multiple antenna system, for optimum combining over all antennas the RAKE receiver becomes two dimensional, with a finger per antenna per multi-path component, as shown in Figure 5-19. The weights now provide maximum ratio combining over all antennas and all multi-path components. This receiver would precede the decoder in an STBC system with multiple receive antennas. For spatial multiplexing there would be a one-dimensional RAKE receiver per receive antenna, and the outputs would feed into the baseband processing in the same way as the antenna outputs in the narrowband system. Hence in either case, an additional RAKE is required per receive antenna, multiplying the complexity of this by n_R .

UNCLASSIFIED

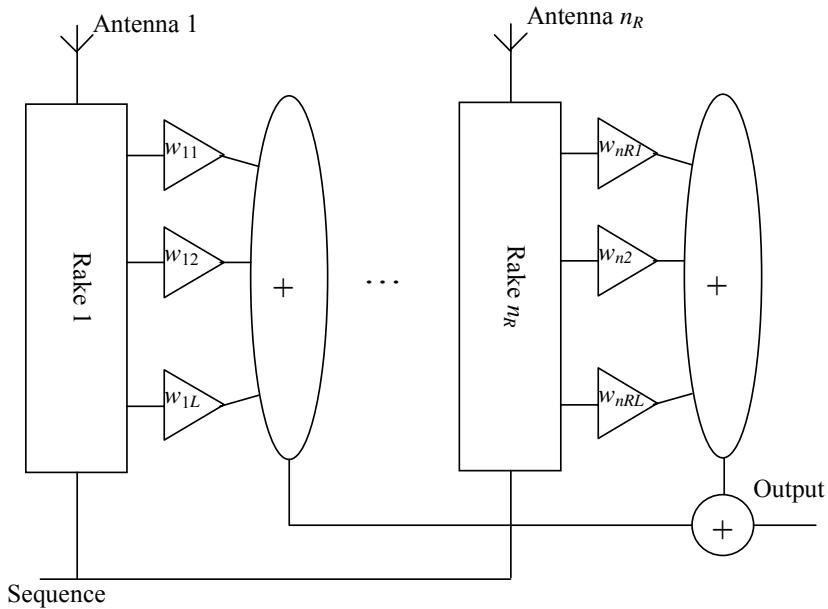


Figure 5-19: Two-dimensional RAKE receiver

Note that in the case of spatial multiplexing CDMA systems, the baseband processing for MIMO reception is very similar to that for multi-user detection. Hence these processes can be combined readily. The effect, and the complexity, is similar to that of $K \times n_T$ single antenna users in the CDMA system, where K is the number of MIMO users.

CDMA systems usually involve forward error correction (FEC) coding to reduce sensitivity to multiple access interference (MAI), and hence to increase capacity. Especially in the case of MIMO-CDMA, the performance of the receiver can be significantly improved by iterative joint decoding, MIMO and multi-user detection. [149] describes such a system, referred to as space-time turbo-coded turbo parallel interference cancellation (ST-TuC-turbo-PIC) for which the receiver is shown in Figure 5-20. In such a receiver, the bit error rate (BER) performance nearly reaches the single-user SISO bound, even with multiple MIMO users, each with data rate increased by the factor n_T , as shown in Figure 5-21.

UNCLASSIFIED

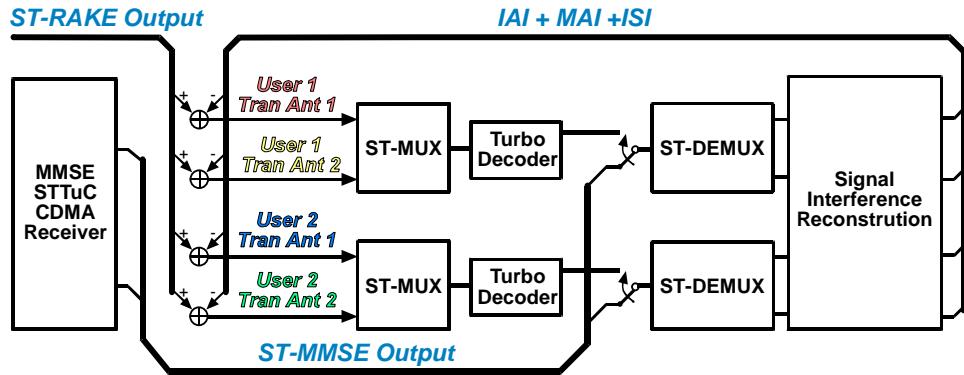


Figure 5-20: ST-TuC-turbo-PIC-receiver

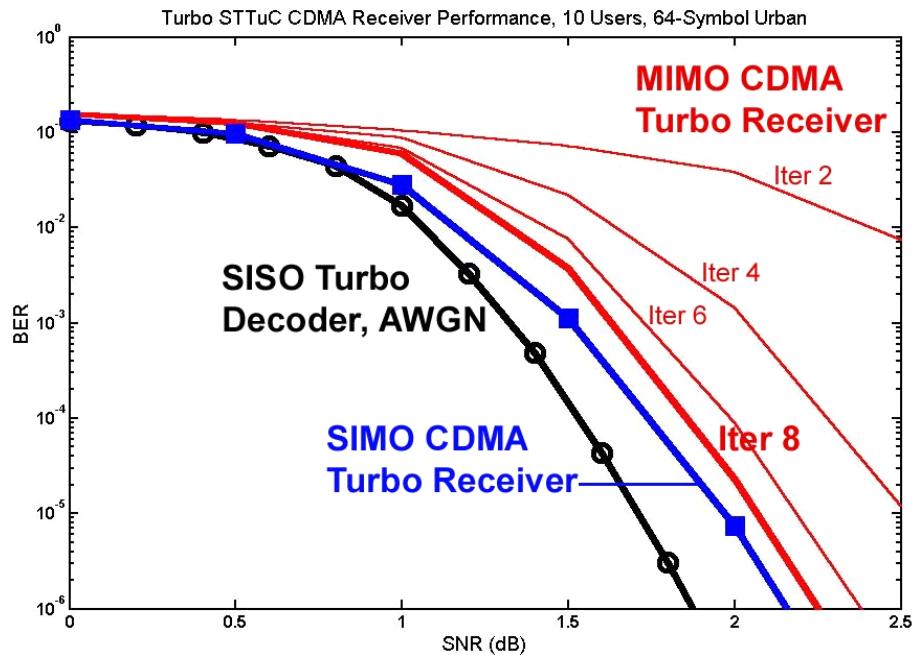


Figure 5-21: Performance of ST-TuC-turbo-PIC receiver

UNCLASSIFIED

The FLOWS project, mentioned above, has also described an implementation of a MIMO-CDMA terminal using spectral-domain processing, and has shown that this can result in a much lower overall complexity [150]. The basic approach involves transformation of the input signal into the frequency domain using a fast Fourier transform (FFT), whereupon all processing, including correlation and the implementation of the RAKE, are performed using software in the frequency domain. This approach is well-suited to SDR implementation since it allows the use of special-purpose FFT processors, which can be implemented very efficiently, and all other operations are implemented in software. (Special purpose Viterbi decoders are also provided for the FEC decoding that is required by most wireless standards.) Figure 5-22 shows the computational complexity of a MIMO-CDMA system implemented in the time domain and in the spectral domain on a general purpose signal processor, showing firstly that such implementation is possible in pure software, and secondly that the spectral domain implementation is much more efficient.

UNCLASSIFIED

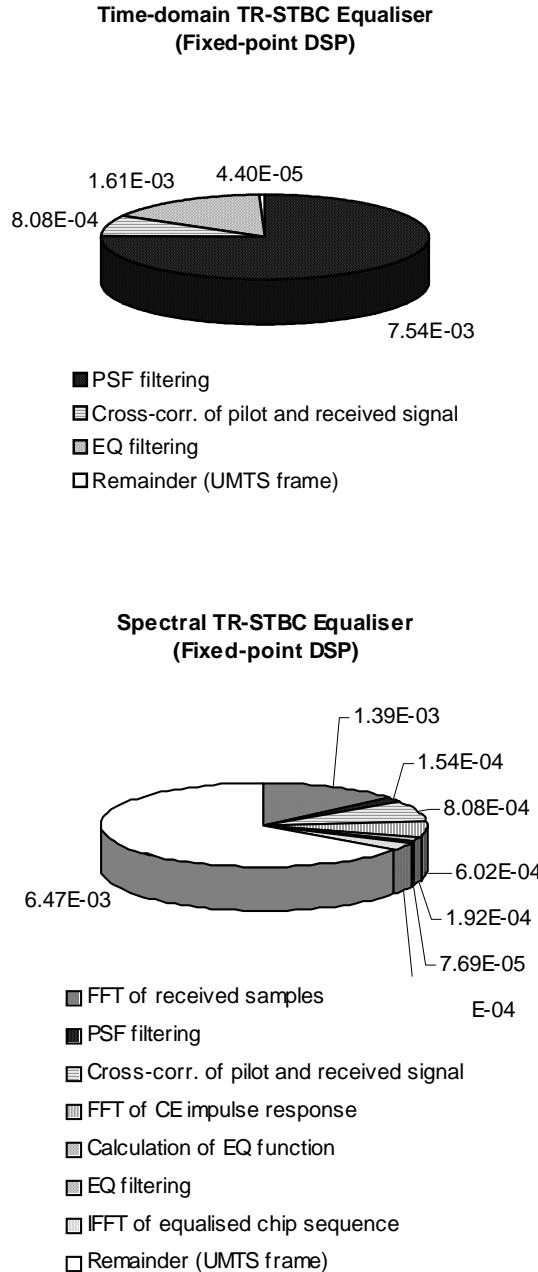
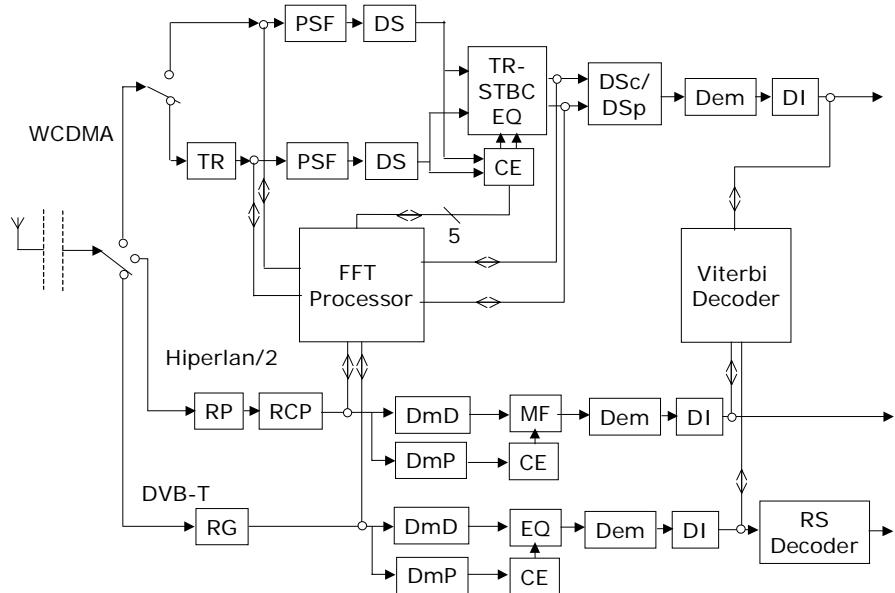


Figure 5-22: Complexity (in terms of processing time) of time domain (top) and spectral domain (bottom) implementation of MIMO-CDMA system implemented on TMS320C64x processor [150]

Figure 5-23 shows a block diagram of a general-purpose baseband processing system based on spectral domain processing. Note that this is also readily applicable to OFDM systems, in which much of the processing takes place in the spectral domain.



CE	channel estimation	DS	down-sample	RCP	remove cyclic prefix
Dem	demodulator	DSc	descramble	RG	remove guard
DI	de-interleave	DSp	despread	RP	remove preamble
DmD	de-map data	EQ	equaliser	RS	Reed-Solomon
DmP	de-map pilot	MF	matched filter	TR	time reversal

Figure 5-23: General purpose spectral domain processor for both CDMA and OFDM systems [150]

5.5.3.2 OFDM

In a multiple antenna OFDM receiver, the OFDM sub-channels are (at least in principle) processed separately by the MIMO baseband processing. This implies that an OFDM demodulator (FFT processor) is required per receive antenna. For example, Figure 5-24 illustrates an OFDM-STBC system. A similar structure would apply to a spatial multiplexing receiver. Hence n_R OFDM demodulators are required, and the complexity of this ‘front-end’ element of the baseband processing is increased by the same factor as in a CDMA receiver. Note that although there is one STBC encoder/decoder per OFDM sub-channel, this does not imply that the complexity of encoding/decoding is multiplied by the number of sub-carriers, since each operates at the OFDM symbol rate.

UNCLASSIFIED

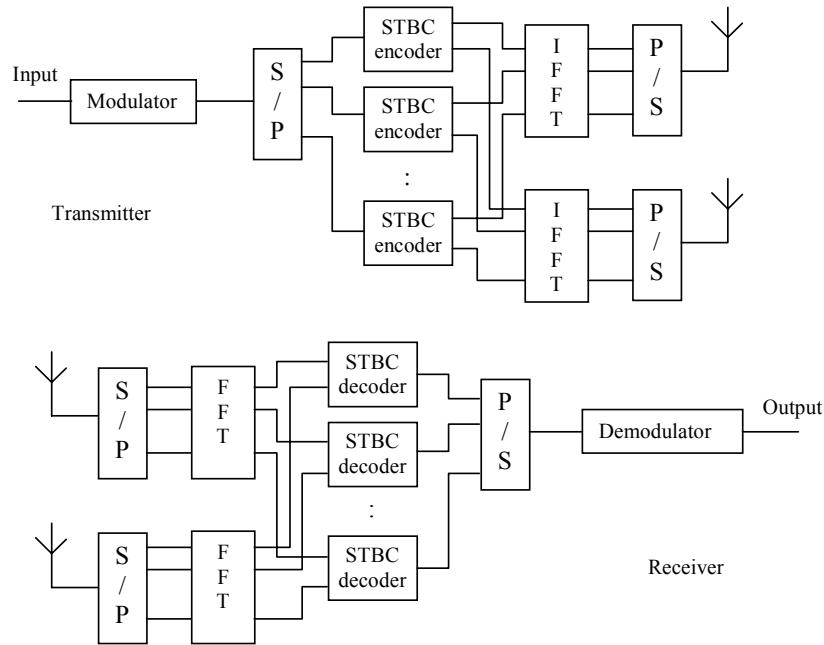


Figure 5-24: OFDM - STBC transmitter/receiver

In OFDM receivers there is an alternative to STBC. The same block code can be applied across the sub-carriers of the OFDM multiplex, rather than in the time dimension, along the sub-carriers. This results in space-frequency block coding (SFBC), which may have advantages on certain channels. This approach is shown in Figure 5-25.

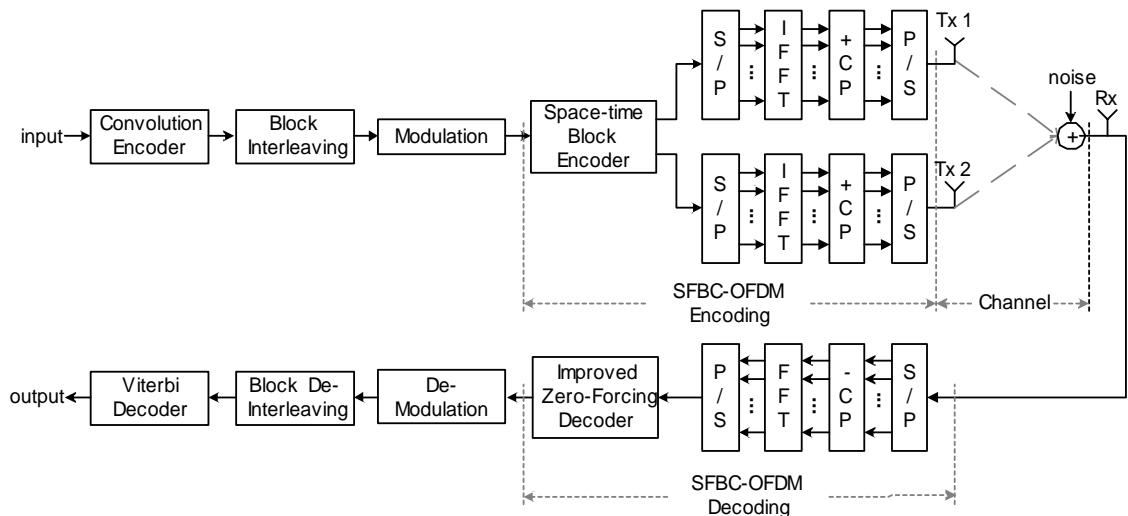


Figure 5-25: Space-frequency block coding

UNCLASSIFIED

5.6

Adaptive MIMO and SDR

As mentioned above, MIMO can benefit significantly from adaptation to the channel, and adaptivity is, of course, a feature of SDR implementations. In addition, since they can potentially adapt to different air interfaces, they can readily adapt to relatively small changes within an air interface. This adaptivity can take various forms, providing various levels of benefit.

Most of these schemes require knowledge of the channel at the transmitter in order to allow adaptation to its characteristics. This must in general be provided by feedback on a reverse channel from the receiver to the transmitter. Such schemes are known as *closed loop*: we have already noted that many of the proposals for future MIMO standards are closed loop. These tend to require that the channels be relatively stationary, since the information cannot generally be fed back for several frame periods of the reverse link, and if the channel changes significantly during that time the information may be outdated and become useless. Also a stationary channel will require a lower overhead for the feedback of channel state information. Some schemes, however, can use information obtained by estimating the reverse link channel at the transmitter, thus not requiring explicit channel state feedback, and some forms of adaptivity need operate only at the receiver, also avoiding feedback.

We will investigate the benefits of some of these forms of adaptivity in this section. To avoid the need to explicitly simulate a large number of potential schemes, which would be beyond the scope of this report, we consider the information theoretic capacity (mutual information) of the systems with various forms of feedback.

5.6.1

Adaptive Modulation and Coding

It has been shown [151], [152] that SISO systems can benefit from the use of adaptive modulation and coding, in which the coding and/or modulation rate is adapted to the current capacity of the channel as it changes due to varying co-channel interference and channel fading. In the absence of this adaptivity, the code rate must be chosen to match the worst-case channel (or that which applies in 90% or more of cases). Adaptivity allows some communication to occur with higher availability, and means that over time the average capacity of the channel can be approached, rather than the capacity of the poorest likely channel.

UNCLASSIFIED

In some MIMO systems, and especially spatial multiplexing, it may in addition be beneficial to adapt the modulation and coding rate of the different spatial sub-channels separately, since otherwise some sub-channels may suffer outage while others are under-utilised. For example, Figure 5-26 shows the capacity distribution of the eight spatial sub-channels corresponding to the eight transmit antennas of an 8×8 spatial multiplexing system, ordered from highest to lowest. The channel has receive antenna elements spaced by 0.2λ , resulting in correlation between elements. MMSE detection is used at the receiver to separate the sub-channels. The figure shows that the median capacity of the sub-channels varies by a factor of four to one. Hence, if the same modulation/code rate is used on all sub-channels, then either the outage capacity of the poorest sub-channel will be high, resulting in a high BER on that sub-channel (and an unacceptable BER overall), or the best channels will be under-utilised by around the factor four. Figure 5-27 shows the difference between the effective capacity distribution of an adaptive system, in which adaptive modulation and coding is used on each sub-channel, and that of a system using the same modulation and coding on all sub-channels. For, say, 10% outage probability the non-adaptive system is limited to around 2.3 bits/s/Hz while the adaptive system can achieve 4.8 bits/s/Hz. Moreover, on average, the adaptive system could achieve the mean capacity, of around 5.5 bits/s/Hz.

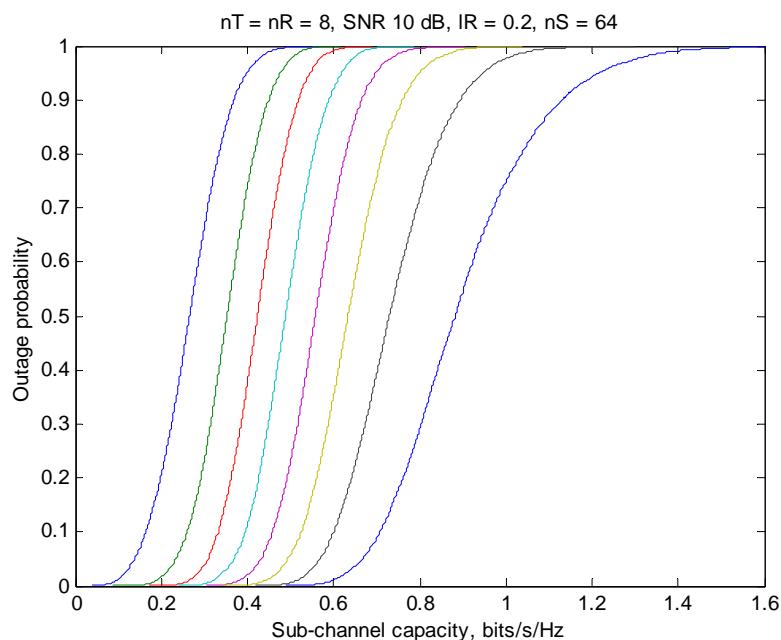


Figure 5-26: Distribution of capacity of sub-channels of 8×8 spatial multiplexing scheme using MMSE reception

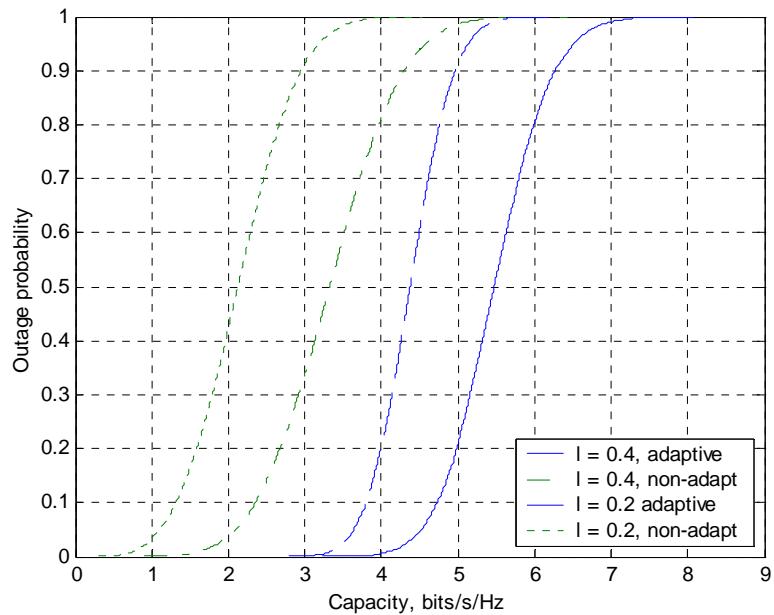


Figure 5-27: Outage capacity for adaptive and non-adaptive 8 x 8 MMSE spatial multiplexing scheme, correlated ($l=0.2$) and uncorrelated ($l=0.4$) channels

5.6.2 Receiver Beamforming

For multiple antenna receivers another advantage is available in systems subject to interference with directional properties, such as cellular systems. This can be applied at the receiver only and does not require feedback. It can readily be shown by analogy with the well-known case of optimum reception in the presence of interference which is non-white in the frequency domain, that there exists an optimum linear receiver for a multiple antenna receiver in the presence of interference which is spatially non-white, i.e., non-uniform in direction of arrival. We refer to this as a *spatially pre-whitening matched filter* and it consists in principle of a matrix which implements a spatially-whitening pre-filter, followed by a maximum ratio combiner (MRC) (effectively a spatial matched filter). The pre-filter is an $(n_R \times n_R)$ matrix which ensures that its outputs are uncorrelated. In effect it exploits the correlation of the interference between the receiver antennas to partially cancel it. Of course, in practice the pre-whitening filter and the MRC will be combined into a single matrix transformation. It turns out to be equivalent to the *max SINR beamformer* which is well-known from the literature of ‘smart antennas’. Another way of understanding its operation is that it attempts to steer nulls in the response of the receive antenna array in the directions of interferers (taking into account their multi-path signals), while also minimising the effect on the wanted signal, so as to maximise the overall SINR.

The optimum pre-whitening filter is given by [153]:

$$\mathbf{W} = \boldsymbol{\Phi}_{\text{int}}^{-1/2} = \sqrt{\boldsymbol{\Lambda}_{\text{int}}} \mathbf{U}_{\text{int}}^H \quad \text{Equation 5-17}$$

UNCLASSIFIED

where Φ_{int} is the correlation matrix of the interference plus noise. The effective channel for the signal then becomes \mathbf{WH} , and the capacity can be found by replacing \mathbf{H} by \mathbf{WH} in Equation 5-2.

Figure 5-28 shows the mean capacity when there is one dominant interferer, signal-to-interference ratio (SIR) 10 dB, and a 4x4 system with eight scatterers. It shows that the pre-whitening filter has a very dramatic effect on capacity at high signal to noise ratio (i.e., where the system is interference limited). Note that the advantage depends on the number of interferers and the number of scatterers, especially on the interfering channel. If the product of these is small, the pre-whitening filter, acting as a beamformer, is able nearly to null the interference, resulting in a potentially large beamforming gain. With a larger number of interferers the interference becomes more spatially white, and the advantage available is much reduced.

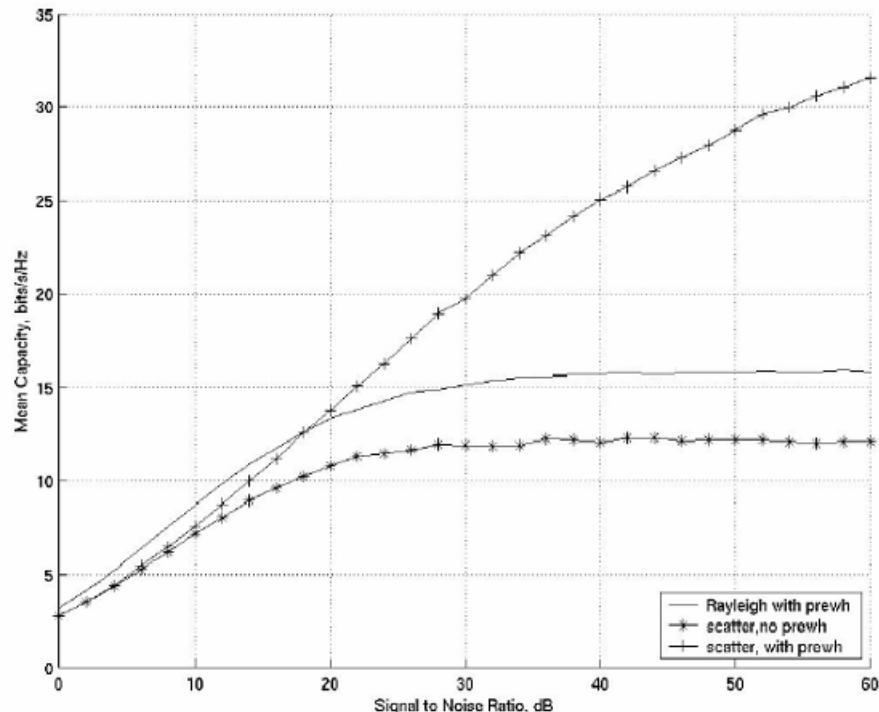


Figure 5-28: Capacity of a 4 x 4 MIMO system with one interferer; a single bounce channel model with eight scatterers ($n_s=8$) and the interference is positioned at 45 degrees to the wanted signal and a SIR of 10 dB is assumed

Figure 5-29 shows the capacity for a simulated cellular system, including the effects of an inverse fourth power propagation-distance law and log-normal shadow fading. Although there is potentially a large number of interferers here, the interference remains sufficiently spatially non-white that there is a capacity improvement of around 50% with the pre-whitening filter.

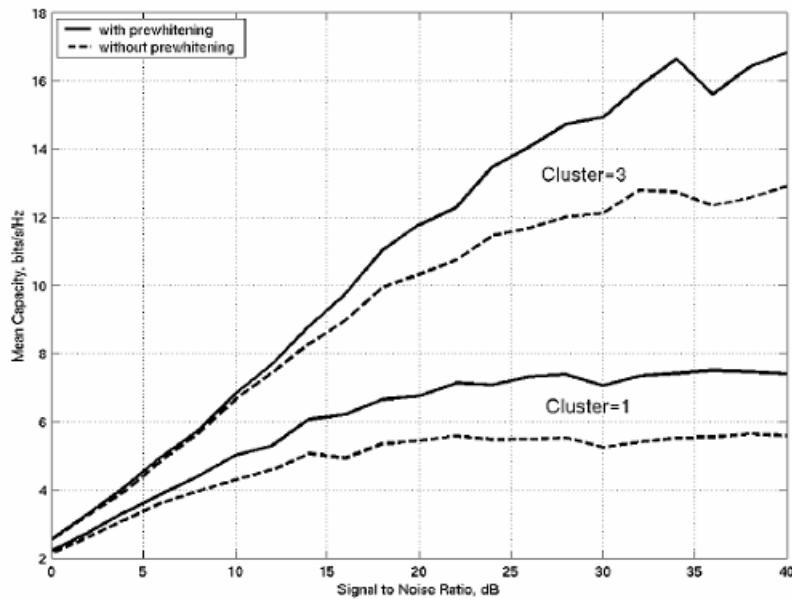


Figure 5-29: Average capacity of 4×4 MIMO cellular system, for re-use factors 1 and $1/3$ (cluster sizes 1 and 3)

5.6.3 MIMO Exploiting Channel Knowledge at the Transmitter

As mentioned above, if the channel is known at the transmitter the transmissions can be adapted to it. This is best understood in terms of the singular value decomposition described in Section 5.2. If the channel is unknown at the transmitter, the best that can be done is to transmit equal power in each channel eigenmode. In fact, conventional space-time codes and spatial multiplexing techniques are designed so that they achieve this, at least on average. However, if the channel is known, and hence also its full singular value decomposition, then the power distribution can be optimised according to Shannon's principle of 'water-pouring' [154].

This principle was originally developed for the case where information is transmitted over a wideband channel with non-white interference or noise. Shannon showed that to maximise the overall capacity for a given total transmit power, the power should be distributed so that the sum of the interference power and the signal power is, as far as possible, constant across the band. This can be visualised by plotting the power spectrum of the interference, and imagining that one pours water into this plot, so that the level of the water is equal across the band. The depth of the water at each frequency then gives the power to be transmitted at that frequency.

UNCLASSIFIED

In a MIMO system, in place of an interference power spectral density continuously variable with frequency, we have a set of discrete sub-channels, and instead of constant gain and variable interference, we have the same noise power in each sub-channel but different channel gains (set by the channel eigenvalues). Figure 5-30 shows how ‘water-pouring’ can be applied in this case. The channel gains can be rescaled by equalisation at the receiver, which results in equal gains but different noise power (b), which then allows distribution of the signal power according to ‘water-pouring’. Note that the result is that the channels with the lowest gains have the least transmit power, and there may be some (in which the rescaled noise power falls above the ‘water line’) in which no power is transmitted at all.

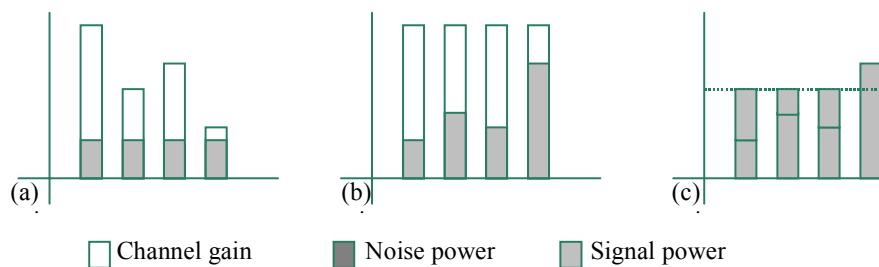


Figure 5-30: Application of 'water-pouring' to adaptive MIMO system: (a) MIMO system with variable channel gains; (b) re-scaling to give constant channel gains, variable noise; (c) application of 'water-pouring' to determine signal power

Mathematically, if the noise per sub-channel is N and the channel eigenvalues are λ_i , $i = 1 \dots n_r$, where n_r is the rank of the channel, then the rescaled noise power in the i^{th} sub-channel is N/λ_i . The ‘water-pouring’ rule then states that the total of noise plus signal power should be as far as possible constant (say P) across all channels, subject to the constraint $\sum_{i=1}^{n_r} S_i = S$. Then the signal power per sub-channel is:

$$S_i = \max\left(P - \frac{N}{\lambda_i}, 0\right) \quad \text{Equation 5-18}$$

where P can be found by solving:

$$S = \sum_{i=1}^{n_r} \max\left(P - \frac{N}{\lambda_i}, 0\right) \quad \text{Equation 5-19}$$

UNCLASSIFIED

In some cases, for some sub-channels, $N/\lambda_i > P$, and hence no power is transmitted in that sub-channel. We will call the number of sub-channels with non-zero transmit power n_a , the number of active sub-channels.

The capacity is then:

$$C = \sum_{i=1}^{n_a} C_i = W \sum_{i=1}^{n_a} \log_2 \left(1 + \lambda_i \frac{S_i}{N} \right) \quad \text{Equation 5-20}$$

where C_i is the capacity of the i^{th} sub-channel.

Of course, to implement this power distribution also requires knowledge of the transmit end eigenvectors, the columns of the matrix \mathbf{U} . The most straightforward implementation of the system then follows Equation 5-4, as shown in Figure 5-31. The data are de-multiplexed into n_a streams, in general with different rates, and encoded and modulated. The resulting signals are then weighted to assign the required transmit power according to Equation 5-18, then applied to the matrix transformation \mathbf{U} , resulting in signals which are fed to the transmit antenna elements. This requires input concerning the channel state at three points: the rate of the encoder/modulators, and hence also the outputs of the de-multiplexer, must be chosen according to the sub-stream capacities given by Equation 5-20; the power allocation, given by Equation 5-18; and the eigenvectors, obtained from the singular value decomposition in Equation 5-3.

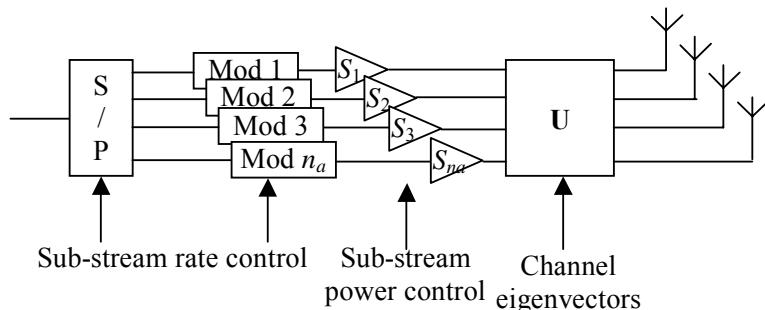


Figure 5-31: Implementation of adaptive MIMO transmitter

Figure 5-32 shows the capacity distribution for an 8×8 MIMO system on an independent Rayleigh fading channel, for signal to noise ratio 0 dB and 10 dB. We note first that the capacity gain is not very large for this case, especially when the signal to noise ratio is large. This is because the channel is full rank and the spread of eigenvalues on a Rayleigh channel is not large. At large signal to noise ratio, P tends to be much greater than N/λ_i for all sub-channels, and hence, according to Equation 5-18, the power is distributed approximately equally among the eigenmodes, as it would be in a non-adaptive system.

UNCLASSIFIED

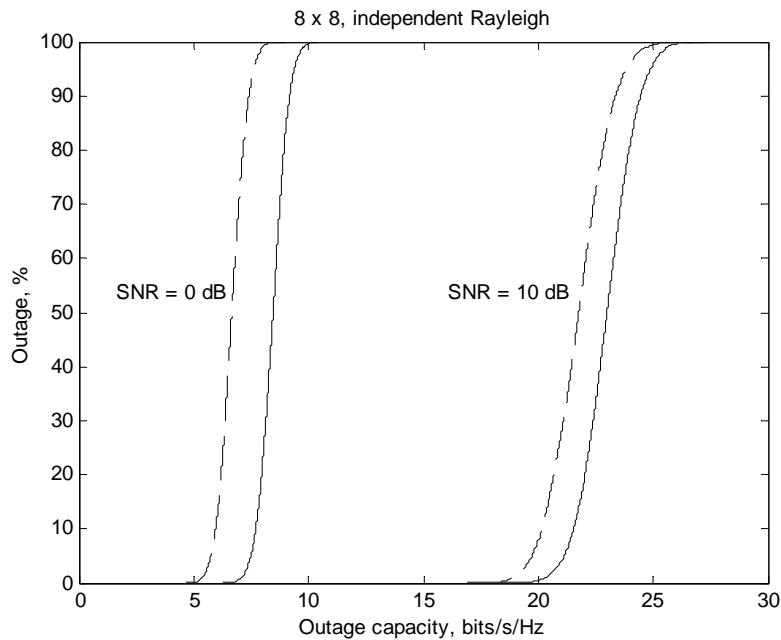


Figure 5-32: Capacity distribution for 8 x 8 MIMO system on independent Rayleigh channel at 0 dB and 10 dB SNR, for adaptive (solid line) and non-adaptive (dashed line) systems

This result suggests that for a symmetrical MIMO system on an independent Rayleigh channel, at least for large signal to noise ratio, the adaptive system has little advantage over the non-adaptive. However, at very low signal to noise ratios, the advantage is large, both in absolute terms and as a proportion of the capacity achieved by the non-adaptive system. In fact, we find that, if the receive antenna array gain is neglected; the average capacity gain of the adaptive over the non-adaptive system is several times that of the MIMO over a SISO system.

The gain is greater in a MIMO system in which the channel matrix is not full rank, that is, the rank is less than n_T , the number of transmit antenna elements. This is because the non-adaptive system must distribute its power equally among the n_T modes that could be generated by n_T transmitting elements, while the adaptive system can concentrate the power in the n_r modes with non-zero eigenvalues. Hence the gain in this case is at least n_T/n_r , even at large signal to noise ratio.

One simple case in which this applies is an asymmetric MIMO system, with $n_T > n_R$. In this case the rank is n_R , and a gain of at least n_T/n_R is available. Figure 5-33 compares the distribution of capacity for an 8x8 MIMO system with 8x4 and 8x2 systems, both adaptive and non-adaptive: whereas the gain is small for the 8x8 adaptive systems, it is much more significant for 8x4 and 8x2 systems.

UNCLASSIFIED

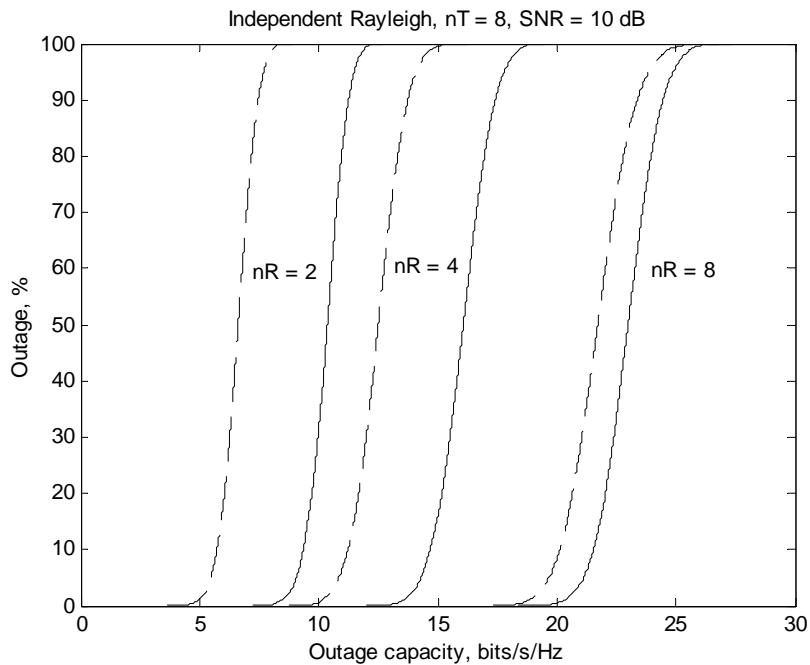


Figure 5-33: Capacity distribution for MIMO system with 8 transmit elements on independent Rayleigh channel at 10 dB SNR, for adaptive (solid line) and non-adaptive (dashed line) systems

Figure 5-34 expresses a similar result in a different way. Here, the number of receive antennas is fixed at four, and the mean capacity for adaptive and non-adaptive systems is plotted against the number of transmit antennas. We note that for the non-adaptive system the capacity tends to a limit with increasing numbers of transmit elements, such that there is little advantage in making the number of transmit elements much greater than the number of receive, but for the adaptive system it continues to increase without limit, such that for highly asymmetric systems capacity may be doubled in the adaptive system.

UNCLASSIFIED

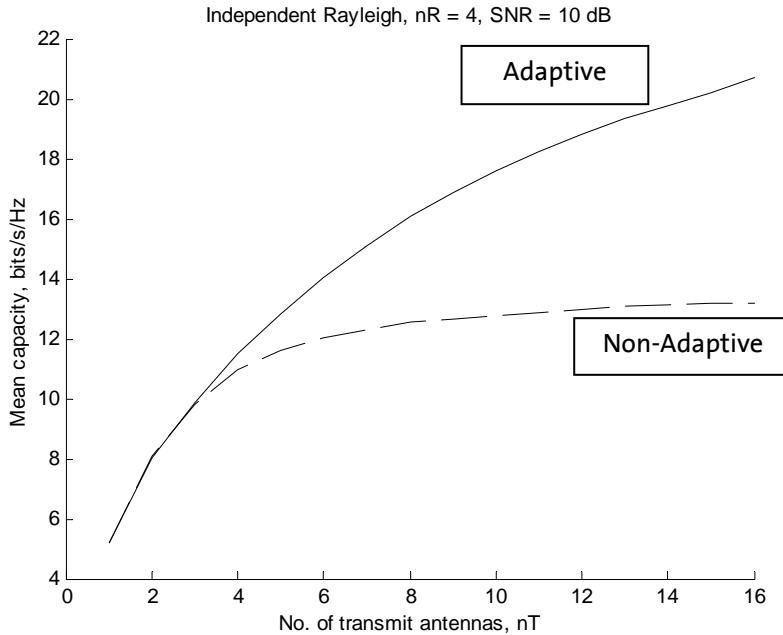


Figure 5-34: Mean capacity of independent Rayleigh channel for four receive antennas on independent Rayleigh channel, SNR 10 dB, for adaptive and non-adaptive systems

We now turn our attention to the finite scatterers channel model, which was introduced in Section 5.2.1 as a more accurate model, in general, than the independent Rayleigh model. Two aspects of this model result in fading which is not independent: firstly, that the number of multi-path components is limited, and secondly, that the spacing of the transmit elements may result in correlation of the fading of adjacent elements.

Considering the first of these, we note that Equation 5-5 implies that the rank of the channel matrix is limited to n_s . Hence, according to the discussion above, we may expect that adaptive transmission will show a gain when $n_T > n_s$. Figure 5-35 shows the capacity distribution for 4×4 , 8×8 and 16×16 adaptive and non-adaptive MIMO systems for eight significant multi-paths ($n_s = 8$). The adaptive system shows small gains for the 4×4 and 8×8 cases, but a more significant gain of around 40% for the 16×16 case.

UNCLASSIFIED

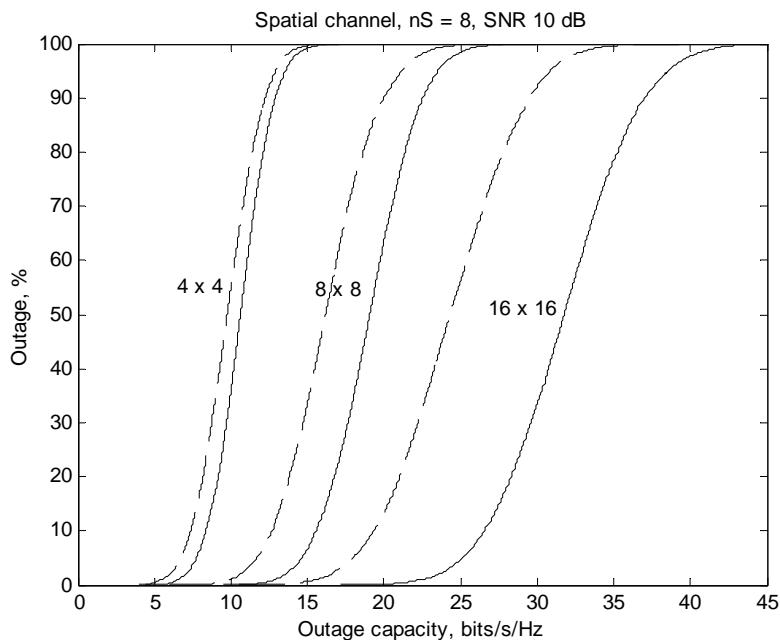


Figure 5-35: Capacity distributions on finite scatters channel with eight significant multi-paths, SNR 10 dB for adaptive (solid line) and non-adaptive (dashed line) systems, and different numbers of transmit and receive antennas

Figure 5-36 gives a different view of the same phenomenon. The mean capacity of symmetric MIMO systems is plotted against the number of transmit/receive elements. In this case, the channel is normalised to remove the effect of the array gain, as discussed in Section 5.4.1 above. As shown in [126], the capacity tends to a limit as the number of transmit/receive elements increases, given by the number of multi-paths. However, the capacity of the adaptive system continues to increase without limit, even without the effects of array gain. Again a doubling of capacity is possible for large numbers of transmit/receive elements.

UNCLASSIFIED

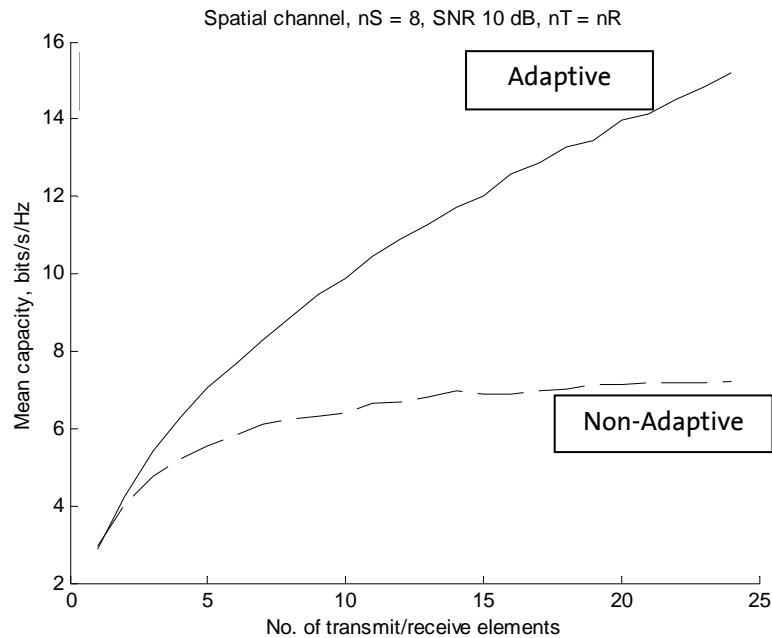


Figure 5-36: Mean capacity of finite scatters channel, normalised with eight significant multi-paths, SNR 10 dB, for adaptive (solid line) and non-adaptive (dashed line) systems, equal nos. of transmit and receive antennas, plotted against the no. of elements

Finally we consider the effect of antenna spacing. Figure 5-37 shows the capacity distribution for adaptive and non-adaptive systems with antenna element spacing from 0.1 wavelengths to 10 wavelengths (at both transmitter and receiver). The number of multi-path components is much larger than the number of transmit/receive elements, so this will not affect the relative performance of the adaptive and non-adaptive systems. Angles of arrival and departure are uniformly distributed over 2π , so that significant correlation begins to occur for element spacing less than about a half wavelength. We observe that for such spacing, the capacity begins to reduce significantly compared to widely spaced elements, but that the adaptive system capacity reduces less rapidly. For a spacing of 0.1 wavelengths the adaptive system has a 50% capacity advantage over the non-adaptive.

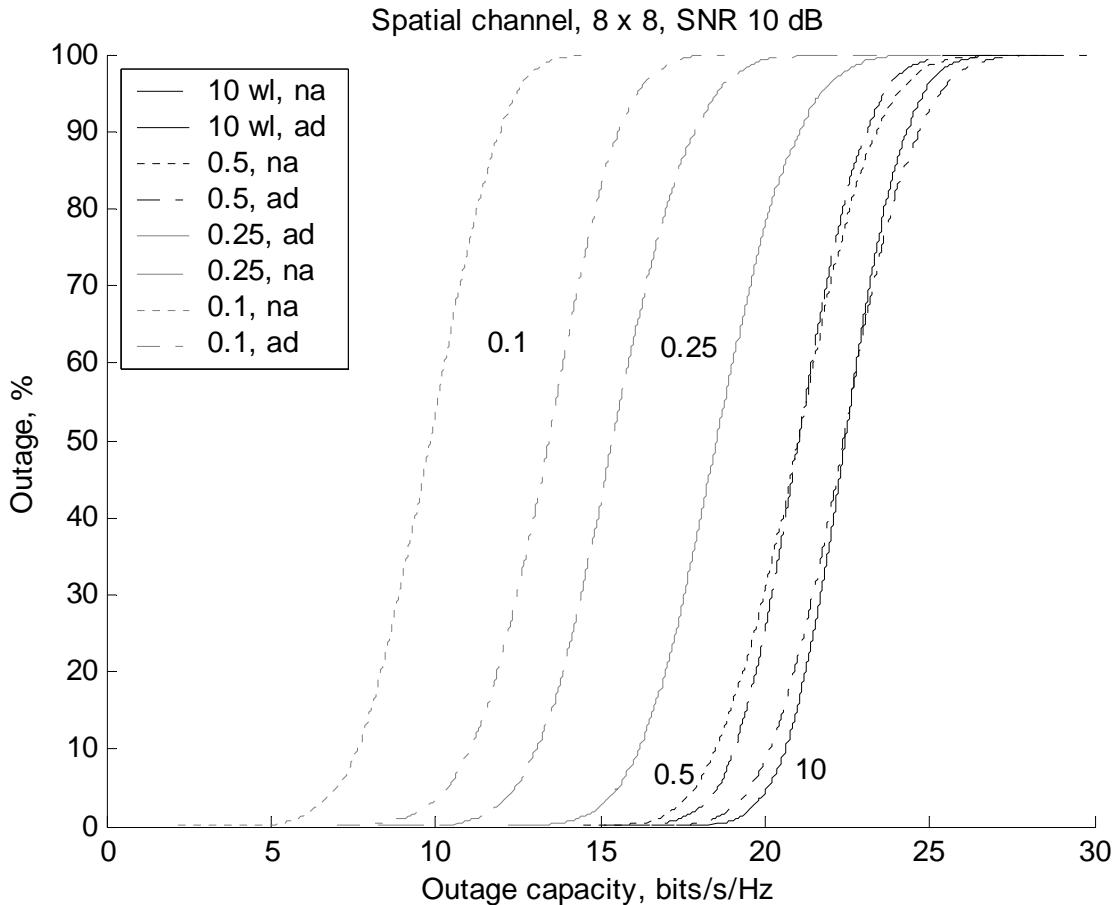


Figure 5-37: Capacity distribution for 8 x 8 MIMO system on a finite scatterers channel with 64 multi-paths components, SNR 10 dB, with different transmit and receive antenna element spacing (in wavelengths) for adaptive (ad) and non-adaptive (na)

5.6.4 On the Issue of MIMO Implementation for SDR

Figure 5-31 has shown the structure of an adaptive MIMO transmitter. The underlying structure of the receiver is essentially the reverse: a matrix transformation, this time using the matrix \mathbf{V}^H , followed by a set of demodulator/decoders, followed by a multiplexer to reassemble the original data stream. Here we consider firstly to what extent adaptation increases the complexity of SDR implementation compared with a non-adaptive system, and secondly how SDR, in comparison with a custom implementation, enables adaptive MIMO. Note that from the point of view of complexity, this adaptive MIMO approach can most readily be compared with a non-adaptive spatial multiplexing scheme, since both are designed to increase capacity.

UNCLASSIFIED

The fundamental complexity per bit of the underlying structure is little different in terms of operations per bit. Although there are multiple encoder/decoders and modulator/demodulators, the total throughput is the same as in the spatial multiplexing system. The matrix transformation at the receiver is equivalent to the MMSE transformation in the spatial multiplexing receiver. The exception is that the de-multiplexing operation at the transmitter is replaced by a matrix multiplication, involving about $n_a \times n_r$ complex multiply-accumulate operations. Moreover, an additional infrastructure of channel estimation, eigensystem decomposition and power control is required, especially at the transmitter. This will give rise to additional complexity, although the computations are in general required only once per frame.

Although this type of adaptive architecture could be implemented using special-purpose hardware such as ASICs, etc., it is clear that it is greatly simplified and its flexibility enhanced by using SDR. SDR will allow the implementation of multiple separate encoder/decoders operating at different rates without requiring duplication of hardware. It is well suited to performing complex matrix multiplications, since specialised processors are available for such functions. Computations required for channel estimation and eigen-decomposition can be implemented efficiently in pure software. It can allow efficiently for variable numbers of active sub-channels, noting that in an adaptive system this may vary widely depending on the channel.

An adaptive MIMO transmitter of the sort illustrated in Figure 5-31 might, however, also give rise to problems in the implementation of the RF chain. In particular, the output of the matrix transformation will consist of the sum of a number of randomly weighted data streams, and thus will have an amplitude distribution resembling noise, with a large peak-to-mean ratio. This may require more linear high power amplifiers. Moreover, perfect separation of the sub-channels at the receiver assumes a linear RF receive chain and high dynamic range in the receiver ADCs. Sub-optimum adaptive MIMO schemes based on antenna selection [155] can achieve throughput very close to the optimum, but reduce the peak to mean ratio of the transmitted signal, since one modulated stream is transmitted on each antenna. These are particularly well-suited to SDR implementation, since it readily allows for a variable number of antennas without requiring redundant hardware.

5.7

Conclusions

We have considered the implications of MIMO for SDR implementation of wireless communication systems from two viewpoints: firstly, the extent to which MIMO creates a ‘barrier’ to the introduction of SDR, and secondly, the extent to which it can be regarded as a ‘complementary’ technology.

The potential barrier arises from the additional complexity that MIMO systems require, and from the fact that they are likely to form an important part of future standards. We have noted from a brief review (in Section 5.3.2) of proposals within 3GPP and Task Group n of the IEEE 802.11 committee that there are a number of MIMO proposals. In particular, the IEEE 802.11n standard is certain to incorporate MIMO.

The additional complexity of MIMO systems arises in both the RF part of the transceiver (including the antennas and the ADC/DAC) and in the baseband processing. These are considered in Sections 5.4 and 5.5 above.

UNCLASSIFIED

The former is irreducible in the sense that MIMO inherently requires multiple antennas and hence also some duplicated RF processing. If a terminal is required to be reconfigured for multiple frequency bands, antenna configuration becomes an issue for MIMO SDR systems, since the optimum antenna spacing varies with the wavelength of the radio signal. Section 5.4.1 shows both that antenna elements can be designed that operate in multiple bands and considers guidelines for the optimum spacing of elements. In cases where there is limited space on a terminal for the antenna array, an approach is proposed consisting of a combination of multi-band and single band elements so as to allow the maximum number of elements to be deployed in each band.

In Section 5.4 we considered the extent to which the RF processing can be shared in a MIMO SDR system. We considered but rejected the use of extensive analogue RF processing to implement MIMO functions. The use of a multiplexed RF chain based on code-division multiplexing, however, is a possibility that might readily be combined with SDR.

The implementation of baseband processing and especially the additional complexity required for MIMO was considered in Section 5.5. Both STBC and spatial multiplexing were considered, these being the most popular MIMO techniques. It was shown that the additional complexity inherent in the basic techniques is quite small and would not constitute a significant barrier to SDR implementation, especially when consider in terms of operations per data bit. Most wireless systems, however, are wideband in the sense of having a signal bandwidth greater than the coherence bandwidth of the channel and require either CDMA or OFDM techniques to be applied to overcome these. These can result in greater complexity in a MIMO system, since, in general, the processing associated with these techniques (FFT or RAKE correlators) have to be duplicated per antenna at both transmitter and receiver. Note that, however, if the effect of MIMO is to increase the data throughput, this still does not necessarily increase the processing required per bit. Moreover, a spectral domain approach is proposed which is well suited to SDR implementation of both MIMO-CDMA and MIMO-OFDM.

Proceeding to the question of whether MIMO and SDR are complimentary (discussed in Section 5.6), we considered the possible advantages of adaptive MIMO systems, since SDR will ease the introduction of adaptivity. We showed that various forms of adaptivity can be deployed to significantly increase the capacity of MIMO systems. Specifically we considered the use of adaptive modulation and coding on the spatial channels provided by a MIMO system and showed that this is necessary to achieve the full information theoretic capacity predicted for MIMO. Adaptive beamforming approaches at the receiver also increase capacity in the presence of spatially non-uniform interference, such as occurs in a cellular system. More significantly, the exploitation of channel knowledge at the transmitter can allow the transmitter to adapt its transmissions accordingly, potentially resulting in large capacity increases. These are greatest in cases where the rank of the MIMO channel is much less than the number of transmit antennas, for example when there are more transmit than receive antenna elements, or when the number of significant multi-path components on the channel is limited.

Overall, we conclude that the introduction of MIMO does not create an insurmountable barrier to the use of SDR techniques in the implementation of wireless communication systems. On the contrary, SDR may enable the introduction of adaptive MIMO techniques and in this sense is a strongly complementary technology.

6 Waveforms



By Stephan Weiss, University of Southampton.

6.1 Introduction

This chapter considers waveforms and is broadly a tutorial and literature review. The promise of achieving interoperability of communication devices between different standards and waveforms has, in the past, been a strong driver behind the development of software defined radio technology. In particular, wireless communications standards have considerably spread out in their number and variety since the surge in the communications market and the desire of various companies and stakeholders to influence the utilisation of the radio spectrum. Although standardisation efforts have been focussed in organisations such as the IEEE, substantial differences remain and continue to exist between different regions, such as Europe, North America and Japan. Interoperability of devices is therefore desirable in order to achieve a level of compatibility between communications equipment both over their lifetime and following possible changes of geographical location.

A further important aspect of interoperability in the context of SDR is that the communications equipment is and remains responsive to various services on offer over the air interface. For example, a communications device might switch from a mobile wireless connection to a WLAN during the download of a file in order to utilise the cheaper of two available modes of operation. Another motivation is provided by the potential savings in hardware, space, servicing and replacement of equipment, and the chance to trigger a gain in spectral efficiency by implementing advanced signal processing algorithms in software.

Therefore, this part of the technology study focuses on waveforms and their potential for integration. Section 6.2 reviews important parameters that characterise waveforms and will have an impact on the ability to integrate them. An overview of various waveforms, as defined in the major wireless standards, is provided in Section 6.3. Modulation methods are reviewed in Section 6.4, which is an important aspect when considering the integration of waveforms, discussed in Section 6.5. A summary is provided in Section 6.6.

6.2 Factors for the Integration of Waveforms

We first explore the characteristics of waveforms, which have an impact on a radio transceiver implementation in general, and an SDR in particular. Here we distinguish between parameters that will affect the front-end and waveform properties that need to be taken into account when considering the baseband realisation and processing.

UNCLASSIFIED

6.2.1 RF Front-End and Up- and Down- Conversion

6.2.1.1 Bandwidth and Band Position

A waveform is usually tied to one or more frequency bands of operation, which is specified by the agency controlling the radio spectrum, and laid down in the standard description. The frequency band is defined by its bandwidth and band. Various waveforms use their assigned frequency bands in different ways, depending on the modulation and multiple access methods. Multiple access schemes share a frequency band between different users in time, frequency or code. Additionally, some systems, like the UMTS terrestrial radio access (UTRA) FDD system and GSM, use separate frequency bands for up- and downlinks, while UTRA TDD, for example, utilises the same frequency bands for both up and downlink transmissions.

In an SDR, the bandwidth and band position have a direct impact on the design and construction of RF front-end, the required sampling rate of the data converters, and the up- and down-conversion between IF and baseband.

6.2.1.2 Sensitivity

The required signal strength for reception of a waveform is referred to as its sensitivity and describes the minimum power level at which a receiver must be able to detect and reconstruct a transmitted message. Sensitivity can vary considerably for different waveforms [92], and is often linked to the data throughput. This is because waveforms offering a high data rate often employ large constellation maps, which are vulnerable to channel noise and therefore require greater SNR for detection. Notice that the same wireless standard might require different sensitivity levels for different data rates. Section 6.3 will highlight several standards where the data rate can be varied over a wide range by utilising diverse constellation maps.

6.2.1.3 Blocker Characteristics

In addition to its sensitivity, a waveform is associated with a so called ‘blocker’ characteristic. In the transmitter, this refers to the power level that may be leaked into adjacent frequency bands, thus setting minimum requirements for the transmit filter and the overall characteristic of the RF front-end by limiting the emitted power outside the band of interest. In the receiver, the blocker characteristic defines the worst case power level of signals in adjacent bands and hence sets out minimum guidelines for the construction of the RF front-end and the receive filters. Blocker characteristics usually are frequency dependent; considering the receiver, the closer frequency bands are to the band of interest, the lower the level of power that will be permitted in that band [92]. This definition is important, since signal components in adjacent bands will leak power into the band of interest, either due to nonlinear behaviour of the front-end, or through insufficiently suppressed aliasing in the sampling stage.

6.2.2 Baseband Processing

6.2.2.1 Modulation Technique

Waveforms employ modulation schemes to map a digital signal onto an analogue waveform. Modulation techniques can be categorised as linear and nonlinear schemes. Linear methods, like QAM are very popular in modern standards, with only a few opting for nonlinear methods such as frequency and phase modulation. An example of this is in Bluetooth wireless technology, where low cost and the potential for a mostly analogue realisation, with high tolerances, is desirable. If modulation methods incorporate spread-spectrum techniques, whereby a symbol is spread over a frequency interval that is much wider than the bandwidth requirements of the information signal itself, this will have an impact on the bandwidth requirement of a waveform.

6.2.2.2 Complexity vs Data Throughput

As mentioned in Section 6.2.1, waveforms belonging to different standards, or even different modes within a single standard, can possess different data rates. Generally speaking, higher data rates imply higher computational requirements on the baseband processing and the low-level functions operating on the bit stream such as channel coders. However, if the symbol rate is increased, then baseband processing such as filtering or modulation increase the complexity further. A doubling of the symbol rate implies that the processing has to be accomplished within half the time interval, whereas increased requirements on filters, e.g., the transmit and receive filters, leads to a doubling in their length. Thus, for a doubling of the symbol rate, the processing cost is quadrupled.

6.2.2.3 Medium Access Control

The frequency band assigned to a particular waveform has, in general, to be shared in some sense, either between up and downlink, as in the UTRA TDD case, or between different users adopting the same waveform for transmission. Therefore, while modulation maps a digital data stream onto an analogue signal, which occupies a certain frequency band with specified signal characteristics, the aim of the medium access controller (MAC) is to manage waveform access in a shared environment. This can take the form of obtaining a timeslot for transmission in a TDD or TDMA scheme, or a specific carrier or code in a FDMA or CDMA scheme. Many standards are based on carrier-sense multiple access with collision avoidance (CSMA/CA), whereby the establishment of a connection between a transmitter and a receiver is based on the receiver identifying a currently unused slot within a multiple access scheme. This slot is then taken up by the transmitter-receiver pair.

While the MAC is often unique for a specific waveform, it is separate from the physical layer (PHY) and has functionality that is located in a higher layer of the protocol stack. Thus the MAC can be expected to be operated in software regardless and is not, therefore, specific to SDR. Therefore, the MAC will not be considered for integration within this study.

UNCLASSIFIED

6.3 Wireless Standards

Waveforms are defined as part of standards for various mobile and wireless communication schemes. In this section we will provide a brief overview of the most important areas of current wireless standards and the associated parameters of the underlying waveforms, e.g., bandwidth, band position, modulation techniques and the target data rates. These parameters decide how the integration of waveforms has to be accomplished, and will form the basis for discussions considering which waveforms can be combined at a reasonable cost and effort in later sections of this study.

6.3.1 Mobile Standards

Several mobile and cordless standards exist for telephony and general communication. These are generally at low data rates, but with low delay requirements in order to enable real-time speech communications. A summary of the most important standards is given below, with a direct comparison provided in Table 6-1 and Table 6-2.

Standard	Application	Modulation	Multiple Access	Bit Rate
GSM	Mobile	GMSK	FDMA/TDMA	270 kbps
GPRS	Mobile	GMSK	FDMA/TDMA	270 kbps
EDGE	Mobile	8-PSK	FDMA/TDMA	810 kbps
UMTS-FDD	Mobile	QPSK	W-CDMA	Variable up to 2 Mbps
UMTS-TDD		BPSK/QPSK	TD-CDMA	
DVB-H	Video	QPSK/OFDM	Broadcast	up to 15 Mbps
		16-QAM/OFDM		
		64-QAM/OFDM		
IS-54/136	Mobile	$\pi/4$ -DQPSK	TDMA	48.6 kbps
IS-95	Mobile	QPSK	CDMA	1228.8 kbps
PDC	Mobile	$\pi/4$ -DQPSK	TDMA	42 kbps
DECT	Digital Cordless Phone	GFSK	TDMA	1152 kbps
PHS	Ditto	$\pi/4$ -DQPSK	TDMA	384 kbps

Table 6-1: Mobile communications standards I [156][157][158][159][160][161]

UNCLASSIFIED

Standard	Area	Spectrum	Ch. BW	Ratification
GSM	Europe	890-915 MHz 935-960 MHz 1710-1785 MHz 1805-1880 MHz	200 kHz	1990
GPRS	Europe	Ditto	200 kHz	1995
EDGE	Europe	Ditto	200 kHz	1999
UMTS-FDD	Europe and most of Asia	1920-1980 MHz 2110-2170 MHz	5 MHz	1999
UMTS-TDD		1900-1920 MHz 2010-2025 MHz		
HSDPA (UMTS)	Ditto	Ditto	5 MHz	2004
DVB-H	Europe	Not defined	5 MHz	
IS-54/IS-136	North America	824-849 MHz 869-894 MHz	30 kHz	1988
IS-95	Ditto	Ditto	1250 kHz	1995
PDC	Japan	810-826 MHz 940-956 MHz 1429-1453 MHz 1477-1501 MHz	25 kHz	1991
DECT	Europe	1881.792-1897.344 MHz	1728 kHz	1988
PHS	Japan	1895-1918 MHz	300 kHz	1995

Table 6-2: Mobile communications standards II [156][157][158][159][160][161]

UNCLASSIFIED

6.3.1.1 GSM

The GSM system commenced operation from 1991 onwards. GSM uses a frequency division scheme between different cells and a TDMA scheme (using a so called burst structure) within a cell to permit communication between a base station and several users. Up- and downlink are operated over different bands in the 900 MHz region. Additional systems operate at 1.8 GHz, often referred to as DCS 1800, and at 1.9 GHz, referred to as GSM 1900 or PCS 1900. Originally only considered for voice communications, GSM has been extended to packet service transmission by means of the general packet radio service (GPRS) protocol.

6.3.1.2 UMTS

UMTS began rollout in 1999 and is considered a 3G technique (to replace the 2G GSM system and '2.5G' extensions). The UTRA protocol is based on spread-spectrum techniques using W CDMA with a 5 MHz bandwidth and offering superior multi-user performance over GSM. UTRA can be operated in either TDD or FDD modes.

6.3.1.3 IS-54, IS-136 and IS-95

IS-54 and IS-136 are North American mobile phone standards operating in the 900 MHz region based on a TDMA scheme. The more recent IS-95 was the first major standard utilising CDMA spread-spectrum techniques. With a narrower bandwidth than UTRA, a new spread spectrum method, CDMA2000, is currently being considered as a replacement for IS-95.

6.3.1.4 Others

Many other mobile standards, such as the digital enhanced cordless telecommunications (DECT) standard, exist. DECT is a flexible digital radio access standard for cordless communications, providing for voice and multimedia traffic. Other future services to mobile handsets will include, for example, digital video broadcast-handheld (DVB-H).

6.3.2 Wireless Local Area Networks

WLANS mostly cover data traffic to and from wireless multimedia devices. WLAN waveforms are mostly covered by the IEEE 802.11 group of standards, which contain a number of protocols operating in different frequency bands and at various data rates. Other WLAN standards include the European HIPERLAN/2 standard as well as the Japanese multimedia mobile access communication (MMAC) system. MMAC operates at various data rates between 20 and 156 Mbps under different mobility scenarios at carrier frequencies between 5 and 60 GHz with bandwidths of up to 2 GHz. MMAC relies on subcarrier modulation and OFDM, similar to HIPERLAN/2. Except for MMAC, the major WLAN waveforms are summarised in Table 6-3 and Table 6-4. Note in particular that IEEE 802.11a occupies the 5 GHz region. This is common with both HIPERLAN/2 and MMAC. Other standards within the IEEE 802.11 group reside in the 2.4 GHz band, whereby local differences exist due to its definition as 2446.5 and 2483.5 MHz in France, 2445.0 and 2475.0 MHz in Spain, 2471.0 and 2497.0 MHz in Japan and 2400.0 and 2483.5 MHz in other countries.

UNCLASSIFIED

Standard	Application	Modulation	Multiple Access	Bit Rate
IEEE 802.11	WLAN	DBPSK/DSS	CSMA/CA	1 Mbps
		DQPSK/DSS		2 Mbps
		2-GFSK	FHSS	1 Mbps
		4-GFSK		2 Mbps
IEEE 802.11a	WLAN	BPSK/OFDM	CSMA/CA	6, 9 Mbps
		QPSK/OFDM		12, 18 Mbps
		16-QAM/OFDM		24, 36 Mbps
		64-QAM/OFDM		48, 54 Mbps
		DQPSK/CCK		11 Mbps
IEEE 802.11b	WLAN	DBPSK/CCK	CSMA/CA	5.5 Mbps
		DQPSK/CCK		11 Mbps
IEEE 802.11g	WLAN	BPSK/OFDM	CSMA/CA	6, 9 Mbps
		QPSK/OFDM		12, 18 Mbps
		16-QAM/OFDM		24, 36 Mbps
		64-QAM/OFDM		48, 54 Mbps
HIPERLAN 2	WLAN	BPSK/OFDM	TDMA	6, 9 Mbps
		QPSK/OFDM		12, 18 Mbps
		16-QAM/OFDM		27, 36 Mbps
		64-QAM/OFDM		54 Mbps

Table 6-3: WLAN standards I [162][163][164][165]

UNCLASSIFIED

Standard	Area	Spectrum	Ch. BW	Ratification
IEEE 802.11	Europe, Japan, North America	2.4 GHz Band	22 MHz	1999
			1 MHz	
IEEE 802.11a	Ditto	5150- 5350 MHz 5725- 5825 MHz	16.6 MHz	1999
IEEE 802.11b	Ditto	2.4 GHz Band	22 MHz	1999
IEEE 802.11g	Ditto	2.4 GHz Band	16.6 MHz	2003
HIPERLAN	Europe	5150- 5350 MHz 5470- 5725 MHz	20 MHz	2000

Table 6-4: WLAN standards II [162][163][164][165]

6.3.3 Wireless Personal Area Networks

For short distance wireless connections, a number of wireless personal area network (WPAN) waveforms exist grouped under the umbrella of IEEE 802.15. These standards cover various versions spanning data rates from 20 kbps to 55 Mbps [166]. IEEE 802.15.4, also referred to as Zigbee, is a low rate WPAN mostly used for low-power sensor networks, which need to operate using battery power over a period of at least two years. Bluetooth wireless technology, used for the wireless connection of personal computer peripherals, for example, is a medium data rate standard, whereas IEEE 802.15.3 contains a number of waveforms providing a very high data throughput. This group of WPAN standards is shown in Table 6-5 and Table 6-6. Most waveforms within IEEE 802.15 are located in the 2.4 GHz band.

UNCLASSIFIED

Standard	Application	Modulation	Multiple Access	Bit Rate
Bluetooth IEEE 802.15.1	WPAN	GFSK (BT=0.5)	FHSS	1 Mbps
IEEE 802.15.3	WPAN	QPSK	CSMA/CA	11 Mbps
		DQPSK		22 Mbps
		16-QAM		33 Mbps
		32-QAM		44 Mbps
		64-QAM		55 Mbps
IEEE 802.15.4	WPAN	BPSK	CSMA/CA	20 kbps
		BPSK		40 kbps
		16-ary O-QPSK		250 kbps

Table 6-5: WPAN standards I [167][168][169]

Standard	Area	Spectrum	Ch. BW	Ratification
Bluetooth IEEE 802.15.1	Europe, Japan, North America	2.4 GHz Band	1 MHz	2000
IEEE 802.15.3	Ditto	2.4 GHz Band	15 MHz	2003
IEEE 802.15.4	Ditto	868.0-868.6 MHz	2 MHz	2003
		902.0-928.0 MHz		
		2.4 GHz Band	5 MHz	

Table 6-6: WPAN standards II [167][168][169]

6.3.4 Wireless Metropolitan Area Networks

The group of IEEE 802.16 standards are known as wireless metropolitan area networks (WMANs). These standards, summarised in Table 6-7 and Table 6-8, cover high data rate wireless communications. The frequency bands allocated to IEEE 802.16 vary dramatically between different countries, spanning frequency ranges between 10 and 66 GHz. The IEEE 802.16 standards are often referred to as WiMAX and are aiming to complement WLAN by connecting hotspots.

UNCLASSIFIED

Standard	Application	Modulation	Multiple Access	Bit Rate
IEEE 802.16	WMAN	QPSK	TDMA	32 Mbps
		16-QAM		64 Mbps
		64-QAM		96 Mbps
		QPSK		40 Mbps
		16-QAM		80 Mbps
		64-QAM		120 Mbps
		QPSK		44.8 Mbps
		64-QAM		134.4 Mbps
IEEE 802.16a	WMAN	BPSK/OFDM	TDMA	Various up to 75 Mbps
		QPSK/OFDM		
		16-QAM/OFDM		
		64-QAM/OFDM		
		265-QAM/OFDM		
HIPERMAN	WMAN	QPSK/OFDM	TDMA	14, 21 Mbps
		16-QAM/OFDM		28, 42 Mbps
		64-QAM/OFDM		56, 63 Mbps

Table 6-7: WMAN standards I [170][171]

Standard	Area	Spectrum	Ch. BW	Ratification
IEEE 802.16	Specific frequency band varies from country to country	10-66 GHz	20 MHz	2001
			25 MHz	
			28 MHz	
IEEE 802.16a	Europe, USA	2-11 GHz	1.5-20 MHz	2002
HIPERMAN	Europe	2-11 GHz	Variable 1.5-28 MHz	2003

Table 6-8: WMAN standards II [170][171]

UNCLASSIFIED

6.4 Wireless Communications

Modulation involves changing one or more features of a signal at the data rate such that the input data can be recovered at the receiver. If signal characteristics can be altered such that M symbols are differentiable at the receiver, then $\log_2(M)$ information bits can be encoded per symbol. Here the term *modulation* is used to refer to digital baseband modulation, not bandpass modulation, where the desired information signal is multiplied with a sinusoid, called a carrier, for more efficient transmission [172]. Several modulation methods, classified as being either linear or nonlinear, are employed in modern wireless telecommunication standards. A selection of modulation methods and multiple access schemes appear in the range of tables from Table 6-1 to Table 6-8 and are discussed briefly in the following sections.

6.4.1 Amplitude Shift Keying

An amplitude-shift keying (ASK) modulated signal comprises of a sequence of real, bipolar symbols, with a finite number of amplitude levels. Hence, the information is encoded in the amplitude of the symbol. ASK is not commonly used in modern wireless standards, but is relevant because of its derivatives.

6.4.2 Phase Shift Keying

PSK is analogous to ASK, except that information is encoded in the signal phase rather than its amplitude. Symbols are a set of complex numbers with uniform magnitude but different phase. The arrangement of the symbols on the complex plane and the codeword for each constellation point may vary [173]. BPSK, QPSK and 8-PSK are used in some WLAN and cellular mobile systems today. Example constellation maps appear in Figure 6-1.

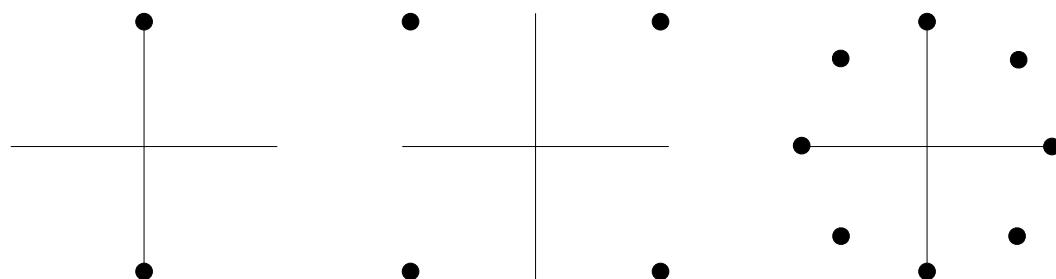


Figure 6-1: BPSK (left), QPSK (middle) and 8-PSK (right) constellation maps

UNCLASSIFIED

Offset QPSK (O-QPSK) evolved from QPSK in an effort to reduce the envelope variation of the transmitted signal and, in particular, avoid symbol transitions that pass through the origin, i.e., zero. Thus, compared to QPSK, O-QPSK minimises the range over which the transmitter power amplifier must remain linear. O-QPSK differs from QPSK because one of the quadrature arms is delayed by half a symbol period so as to eliminate the possibility of both arms being zero at the same time [174]. Another variant of QPSK that aims to reduce envelope variation and avoid zero crossings is $\pi/4$ QPSK. Here, two QPSK constellations are used, one rotated by $\pi/4$ with respect to the other. The transmitter alternates between the two constellations at the symbol rate. Thus, compared to QPSK, the maximum phase shift between symbols is $\pm 135^\circ$ compared to 180° .

Another common modification to PSK is differential PSK (DPSK) and, in particular, differential BPSK (DBPSK), differential QPSK (DQPSK) and differential $\pi/4$ -QPSK ($\pi/4$ DQPSK). These systems encode bits of information in the phase *difference* between successive symbols [172]. That is to say, each constellation point specifies an incremental phase rotation to be applied to the preceding transmit symbol in order to determine the current one. In this way, phase coherence between the transmitter and receiver is not required because the reference phase is taken from the preceding symbol.

6.4.3 Quadrature Amplitude Modulation

In PSK the amplitude of the transmitted signal is constrained to remain constant, thereby yielding circular constellations exemplified in Figure 6-1. By allowing the amplitude of the symbols to vary as well, a new modulation technique, referred to as QAM, is formed. Several variants of QAM exist, with different constellation patterns and number of points [173]. M -QAM is a common notation, where M is the number of legitimate symbols. Hence, QAM is a cross between ASK and PSK, which is evident from symbol map for square 16-QAM, shown in Figure 6-2.

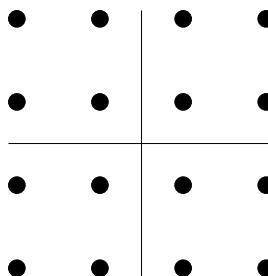


Figure 6-2: Square 16 - QAM constellation map

UNCLASSIFIED

6.4.4 Orthogonal Frequency Division Multiplexing

OFDM applies the principle of multi-carrier (MC) processing by dividing the symbol stream s_n , into P symbol sub-streams, each with a lower rate and requiring a smaller transmission bandwidth. Conceptually, each symbol sub-stream modulates a separate carrier, as shown in Figure 6-3. In practice, however, this is achieved more efficiently by subjecting the symbol stream s_n to an inverse fast Fourier transform (IFFT) and using the result to modulate a single carrier [175]. The reverse process occurs at the receiver. Wireless standards that employ OFDM may differ in the number of IFFT points P , while the symbols s_n may be derived from data bits via a variety of modulation schemes that include BPSK, QPSK, and M-QAM [174].

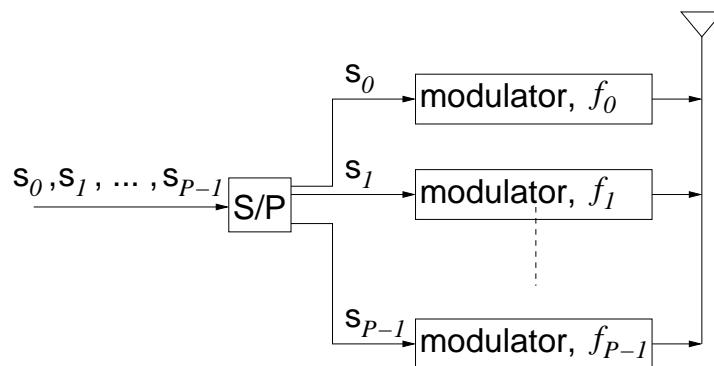


Figure 6-3: Conceptual diagram of OFDM modulation

6.4.5 Direct Sequence Spread Spectrum

Direct sequence spread spectrum (DSSS) or direct sequence CDMA (DS-CDMA) is a processing scheme whereby each symbol s_n in a sequence is multiplied by a spreading code c_k to derive a higher-rate, higher-bandwidth transmit sequence u_k [176][177]. Figure 6-4 portrays an example of this procedure. Each transceiver has its own pseudo-random spreading code which is approximately orthogonal to all others. To recover the symbols s_n , the receiver performs a time correlation once per symbol period between the received signal and the transmitter's spreading sequence. Different types of spreading codes with different properties and lengths are utilised [178] and BPSK, QPSK and QAM are popularly used to obtain the modulating symbol sequence s_n .

UNCLASSIFIED

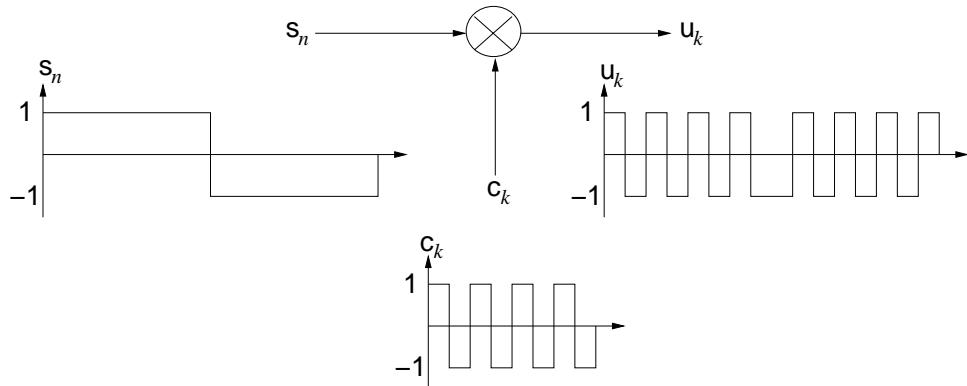


Figure 6-4: DSSS modulation

6.4.6 Frequency Hopping Spread Spectrum

Frequency hopping spread spectrum (FHSS) is a spread spectrum technique in which modulated data bursts are transmitted with pseudo-random carrier frequencies. The instantaneous bandwidth of the transmitted signal is relatively narrowband in comparison to the bandwidth over which the hopping occurs [177][179]. Gaussian frequency-shift keying (GFSK) (see Section 6.4.7) is used to obtain the narrowband signal in a number of WLAN and WPAN systems. Hopping frequencies, instantaneous bandwidth, maximum and minimum burst length and the pseudo-random hopping sequence generator will vary from standard to standard.

6.4.7 Gaussian Frequency Shift Keying

GFSK is a derivative of frequency-shift keying (FSK). FSK entails that the data sequence is first ASK modulated, oversampled by a factor N , scaled by $2\pi h$, where h is the modulation index, and then used to drive a voltage-controlled oscillator (VCO). In this way, a finite number of transmitted signal frequencies, equal to the order M of the modulating ASK signal, are possible [179]. GFSK differs from FSK because the oversampled ASK signal s_k is passed through a Gaussian filter $g[k]$ before proceeding to the VCO. The effect of the Gaussian filter is to reduce bandwidth occupancy. Figure 6-5 shows the signal flow for GFSK modulation. Important parameters that need to be specified include the bandwidth-time product (BT) of the Gaussian filter and the modulation index h . GMSK is an important form of GFSK in which $h=0.5$.

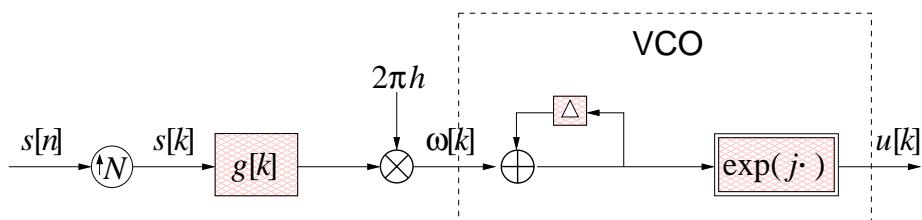


Figure 6-5: GFSK modulation

6.5 Efforts at Integrating some Wireless Standards

Having considered the major modern wireless waveforms and some of their parameters, we will next review efforts geared towards their realisation in software and their integration. Here, the challenges are in the analogue front-end, where the parameters reviewed in Section 6.2.1 (i.e., bandwidth, band position, sensitivity and blocker characteristics) have an impact on the implementation, are likely to be different from those to be found in baseband implementations. In baseband implementations, other integration parameters are likely to influence the design, as reviewed in Section 6.2.2.

We here distinguish between RF and IF front-end integration for single and multi-band band operation. In a single band system, the system operates over a fixed frequency range and may integrate one or more waveforms that reside within the specified band of operation. In contrast, multi-band radios possess a front-end that spans several different frequency bands. In the latter case, the level of integration is of interest. Multi-band radios can, in principle, be based on completely separate hardware and a common baseband processor. A higher level of integration for multi-band radios is accomplished in so called multi-mode implementations, whereby as much as possible of the analogue RF front-end is reused when switching between the frequency bands of operation.

In the following sections a collection of works on the topic of waveform integration is reviewed. The presented literature is reviewed with respect to their extent and functionality in combining, say, wireless mobile networks with other wireless standards.

6.5.1 Single Band Systems for Mobile Standards

Systems relying on the realisation of a single signal path in the front-end may not necessarily be driven by the desire to integrate multiple waveforms. Instead, the motivation might be to implement a receiver with functionality that is superior over an analogue or digital baseband solution.

Experimental SDR systems that have been reported in literature include, for example, the UMTS UTRA TDD mode implementation [100]. This system is aimed at receivers with multiple antennas and uses 16 C6000 baseband DSPs and dedicated down-conversion hardware. This system targeted base station applications and, despite the processing power, is limited in the number of users due to processing constraints. According to [100], more than 80% of the computational resources are required for matched filtering and detection of the signals with just three users. The chosen approach is scalable and, by increasing the computational resources, more users could be supported.

UNCLASSIFIED

Other experimental systems for UMTS include a single chip solution for a direct conversion receiver for UMTS using W CDMA [180]. The analogue band selection filters are tuneable for the various channels within the UMTS standard by means of weighted switched-capacitor matrices. Thus, both the bandwidth and the frequency of a channel in the uplink or downlink can be adjusted. The receiver described in [180] operates at up to 16.384 million chips per second (Mcps). Due to the high sensitivity of the UMTS waveform and the fast clock rate of the system, substrate coupling forced the design to be spread over a total of 4 different dies. Similar systems have been reported some years earlier for ISM-band test systems in the range of 902 to 928 MHz using direct sequence [181] and frequency hopping spread spectrum techniques [182]. The latter techniques also contained substantial processing in the baseband, but their receivers did not have to rely on substantial band selection in the analogue domain prior to conversion.

6.5.2 Systems Integrating 2G and 3G Wireless Mobile Standards

The desire to create mobile communications equipment that is compatible between different standards has been strong. This includes the compatibility between the established GSM standard and the more recent UMTS standard in Europe, and between standards in different geographical regions, such as the integration of the North American IS-95 standard with the European UMTS standard. Consequently, a large number of researchers have proposed solutions and presented test systems for multi-band radios based on the interoperability between second and third generation systems and across the globe. The targeted waveforms of these standards occupy the frequency band between 800 and 2200 MHz. Antennas suitable for operation over such a working range have been presented in [183], for example. Here, we will only focus on the antenna processing, down-conversion and baseband processing.

UNCLASSIFIED

6.5.2.1 Theoretical Studies

General conceptual systems include the study in [184] and [185], where mobile standards in the 2 GHz region were considered for integration. This system aimed to combine the DCS 1800 and PCS 1900 GSM waveforms, DECT and UMTS in the frequency band spanning 1710 and 2170 MHz, a bandwidth of approximately 500 MHz. At the time of the study, analogue-to-digital conversion techniques for operation over such bandwidths were not available and potential solutions based on a direct conversion approach suggested the future use in base station receivers only. As will be seen in Section 6.5.2, such problems can be bypassed in more recent SDR implementations targeting 2G and 3G integration.

Data acquisition and down-conversion methods for the combination of various mobile wireless standards using multi-band SDRs have been addressed by a number of researchers. Integration efforts in this part of the radio have been aimed at various combinations of GSM, UMTS, and IS-95 [186][187], UMTS and EDGE [188], GSM and UMTS [189][190], GSM and PDC [191], GSM, EDGE, and UMTS [192][193][194][195], GSM, DCS, and UMTS [184] as well as UMTS and CDMA2000 for IS-95 [196].

Concerning the baseband processing, substantial contributions have been made by Jondral's group in Karlsruhe on the parameterisation of waveforms [197][198][199][200][201][202]. Rather than reloading entire software modules, a common processing structure is found, whereby parameter changes result in the adoption of a different waveform for processing. Concerning the integration of modulation techniques, linear schemes such as TDMA, FDMA or CDMA approaches can be represented in a common linear algebraic framework. Beyond such linear schemes, it is well known that, although a nonlinear phase modulation, GMSK, as used in GSM, can be expressed as a near-linear scheme [203]. Even the GFSK modulation applied in Bluetooth wireless technology can be linked to a linear partial response modulation method [204].

The general approach for the parameterisation of the baseband modulation and detection in [197][198][199][201] comprises of a precoding stage, a symbol mapping and non-return to zero module, the quadrature demodulation and a stage employing finite impulse response filters for matched filtering and sequence detection. Not every waveform uses all of these functional blocks, but composition allows the designer to implement the mobile wireless standard waveforms that had been anticipated in the conception of the parameterisation approach.

Specifically, while initial work was limited to 2G systems [197][198], this has been extended to include the 3G case in [199][200][201] by incorporating the latter standards into the parameterisation of modulation methods. The integration of 2G and 3G systems by parameterisation requires an approximation of GMSK, as used in GSM, for integration with QPSK symbol encoding, as employed in UMTS. The authors present simulations showing that no significant degradation in performance is incurred in this approximation.

UNCLASSIFIED

The medium access control component, using CDMA for UMTS and FDMA/TDMA for GSM, is not addressed, but equalisation is implemented for all systems. A similar idea is discussed in [199], addressing the baseband modulation stages of the three standards GSM, DECT and UMTS. A wider range of international cellular and cordless standards is discussed in [201], presenting a complete parameterisation of the modulation stage. The authors in general do not discuss the complexity, and rely on the scalability of a hypothetical experimental system in order to match the computational requirements. Baseband implementation of GSM and CDMA based systems is also reported in [205], where both modes are aligned in terms of pulse shaping and symbol mapping.

Efforts similar to the baseband modulation stage have been made in Jondral's group with respect to the parameterisation of channel coders [202], as well as coding [199] and equalisation schemes [206].

6.5.2.2 Implementation of Experimental Systems

Test systems implementing mobile standard integration have been reported. Examples include [207], [208] and [209]. A single chip solution is presented in [207] for combining the European GSM and UMTS standards. This approach is extended in [208] by expanding the dual-mode SDR implementation to four channels, additionally incorporating DCS 1800 and PCS 1900. These solutions are based on analogue RF and IF filters that can be tuned via switched-capacitor matrices, such that a single signal path leads from the RF front-end to the data converter operating at IF. Only the low-noise amplifier is designed and implemented separately for each of the integrated frequency bands. The system in [208] can, therefore, be regarded as a prototype for a multi-band, multi-mode implementation, whereas older realisations, such as [210], for example, necessitate the implementation of a separate down converter for every frequency band to be covered. Similar approaches are followed in [211][212][196], where W-CDMA and IS-95 are combined for operation in base stations. The processing in [196][211][212] relies on different analogue front-ends for each standard, but a common ADC. This is followed by down-conversion stages for each of the various bands.

In [101], GSM and UMTS are combined in an SDR test-bed comprising a FPGA for high-speed front-end processing and a C6000 DSP for baseband processing. For both GSM and UMTS, the down-conversion is realised on the FPGA. GSM requires approximately 80% of the CLBs taken up by UMTS. Baseband processing is performed on a DSP in the case of GSM. For UMTS, however, substantial portions have to be run on the FPGA due to its high data rate. The system is reconfigurable and the device's operational mode can be altered from GSM to UMTS and vice versa by means of software downloads.

In parallel to the integration of civil waveforms, the harmonisation of military communications equipment on a software basis has been a target for a good number of years. For example, the SpeakEasy system aims to combine more than 15 different US military radio standards and waveforms [213][214][215]. Other defence programmes for the integration of military communications equipment have more recently focussed on the waveform description language (WDL) [216], which is a high-level software description of waveform building blocks. With the use of the latter, waveforms can be assembled in a process not unlike block based simulation languages.

UNCLASSIFIED

6.5.3 Systems Integrating 3G and WLAN

6.5.3.1 Theoretical Studies

A conceptual quad-band system incorporating the 5 GHz band of HIPERLAN, in addition to the various European mobile communications frequency bands, is discussed in [92]. Different RF filter banks are used to access the four groups of frequency bands that are located closely enough for common RF processing. These include (i) GSM, (ii) DCS, PCS, DECT and UMTS, (iii) Bluetooth wireless technology and (iv) HIPERLAN/2. A similar integration is demonstrated in [183] in order to operate the GSM, DCS, PCS, UMTS, HIPERLAN/2 and IEEE 802.11a bands, although the problem is only highlighted in terms of a suitable antenna design, such that a single antenna operating across the desired range of frequencies is possible. Further, in [217], UMTS/W CDMA are combined with GSM/GPRS and IEEE 802.11b, Bluetooth wireless technology and GPS on a DSP platform. Here, UMTS is generally considered the more computationally complex task, despite its smaller data throughput as compared to IEEE 802.11b.

A generic solution integrating UMTS using W CDMA and WLAN standards such as HIPERLAN/2 and IEEE 802.11a based on OFDM is discussed in [218]. The paper reflects on a consensus in industry with respect to general computational complexity of the various systems, which is reproduced in Table 6-9.

Standard	Complexity
GSM	10
GPRS	100
EDGE	1000
WLAN	5000
UMTS/W-CDMA	10,000

Table 6-9: Approximate complexity of various waveform applications [218]

6.5.3.2 Implementation of Experimental Systems

The implementation of a tri-band system operating the Japanese mobile and WLAN standards is discussed in [219]. The front-end downconverts three frequency bands at 1.5, 1.9 and 2.45 GHz. A number of analogue amplifier stages are jointly utilised by the different waveform manifestations and the filters can be switched according to the current mode of operation. Software is executed on a DSP chip and can be reloaded over the air interface if the operation mode is to be changed.

Single chip realisations for integration spanning a wider frequency range are presented in [220], where two separate RF paths are employed to directly convert both UMTS using W CDMA and WLAN standards in the 5.8 GHz band down to baseband. Similar efforts to combine mobile communications standards with waveforms in the 5 GHz region are discussed in [221], integrating UTRA FDD and HIPERLAN/2, in [222], combining IS-95 with WLANs, and in [219], [223] and [224], where PHS and IEEE 802.11 are integrated.

UNCLASSIFIED

6.5.4 WLAN and WPAN Integration

Wireless LANs can be integrated relatively simply if operated in the same frequency region. An example of this are Bluetooth wireless technology and IEEE 802.11b [225][226], situated in the 2.4 GHz band. The high specification of the computing platform due to the requirements of it supporting the WLAN standard can be exploited to incorporate a standard of lower cost. In many cases, this leaves headroom for enhancements in the implementation of the low cost standard, e.g., Bluetooth wireless technology [227][228][229].

Bluetooth wireless technology and HIPERLAN/2 integrated solutions have been investigated and implemented in [230], [231], [232], [233], [234] and [235]. While the test-bed operates in real time [235], the front-end is defined by two separate paths for the 2.4 GHz Bluetooth signal and the 5 GHz HIPERLAN/2 signal. Again, Bluetooth wireless technology is an algorithm of negligible cost when compared to the WLAN system, triggering the potential for improving the Bluetooth transmission. However, in an effort to contain the complexity of the baseband implementation, no channel or Viterbi decoding could be implemented for the WLAN realisation due to a lack of processing power.

An integrated radio system combining the WLAN 802.11a/b/g standards operating at 2.4 and 5 GHz is reported in [236] for both receiver and transmitter. Two separate RF paths are implemented for the 2.4 GHz and 5 GHz signals, but the ADC is shared. The baseband processing for the maximally 54 Mbps WLAN standards is implemented using an ASIC rather than a reprogrammable DSP in order to cope with the computational complexity of the system.

6.5.5 Further Integration Issues

In addition to the parameterisation approach discussed in Section 6.5.2, researchers have pursued the investigation of general modulation schemes. Since OFDM and CDMA offer many advantages for multiple access, high data rate and the ability to cope with inter-symbol interference over a broadband channel, schemes have been created which can mimic both OFDM and CDMA systems by parameter selection. This permits the radio to switch between MC CDMA, MC DS CDMA and MT CDMA, which are being considered for future wireless standards [237], [238], [239] and [240]. Further, a frequency hopped multi carrier CDMA scheme has been developed in [241] with particular attention to the integration of IS-95 and W CDMA.

In a combination of mobile, WLAN and WMAN baseband functionality, a common FFT/IFFT structure has been utilised to implement a frequency domain equalisation scheme for both CDMA and OFDM systems [242]. Equalisation is particularly important in communication systems operating at high symbol rates. Interestingly, the work in [242] is one of the few that considers an SDR integration incorporating a WMAN standard.

6.6

Conclusions

In this chapter we have reviewed various parameters associated with waveform implementations that have a direct impact on the ability for integration into an SDR. A tutorial and literature review of a range of wireless standards, with some analysis has been presented. A distinction has been made between the RF front-end and the down-conversion stages, and the baseband implementation. The implementation of the RF front-end and the down-conversion stages is governed mostly by the bandwidth and band position of a waveform, as well as its sensitivity and blocker characteristics. The baseband implementation, on the other hand, is influenced by the modulation mode and the complexity of a waveform, which is often related to the data throughput afforded by the specific standard. Subsequently, the major wireless waveforms, their modulation techniques, and integration efforts have been summarised.

Considering integration efforts reported in the literature, it is noted that RF front-end implementations of waveforms have reached considerable complexity, with quad-band systems in experimental single chip fabrication offering support for a wide range of standards. Such systems are made possible through the availability of flexible antennas for wide frequency ranges. (Wideband antennas are discussed further in Chapter 2 and in [183].)

RF analogue hardware is generally tied to a particular frequency band. However, techniques for tuning and switching have been applied in [180] and [208], for example, thus only requiring a single ADC for several frequency bands. We note that, however, decisions as to which frequency bands can be tuned to must be made at the design stage.

As discussed in Section 6.5, integration has matured for the combination of 2G and 3G wireless standards, with single chip solutions covering the frequency range up to 2.2 GHz for UMTS. The most costly component in SDR currently is the baseband UMTS implementation, which exceeds even WLAN standards due to the need for advanced and complex signal processing algorithms designed to mitigate the effects of fading in a highly mobile environment.

The integration of higher frequency bands, such as the 5 GHz regions of HIPERLAN/2 and IEEE 802.11a, have been considered in theory and suggested solutions for the RF front-end and the baseband processing can be found in the literature. However, experimental systems that are fully operational in an SDR sense have not yet been reported. The realisation in [236] is an ASIC design, while [235] have omitted the high complexity channel decoders from their implementation. Finally, the IEEE 802.11b implementation in [224] is not fully supported unless extra computational resources can be added. Thus, in all these implementations, the lack of sufficient baseband processing power is currently the limiting factor.

Very little evidence has been found so far for the integration of WMAN standards, which may be due to their high complexity and their high geographical variations in band position.

UNCLASSIFIED

Given the survey of waveforms and their integration in this study, it can be expected that a rise in processing power for baseband components might stimulate the full integration of mobile and WLAN standards. This would be particularly useful for mobile handsets or handheld multimedia devices. Thus, for example, lower cost connections with higher data rate services could be invoked if the user were within range of a WLAN. Furthermore, compatibility with international standards such as IS-95 is a useful and realistic feature for handsets.

The creation of future-proof devices may be limited to baseband processing since the analogue RF components might be tuneable, but only within the range of bands considered at the design stage. Therefore, if new standards are assigned different frequency bands that are not considered in the initial SDR hardware design, the extension or reconfiguration of the system would be difficult if not impossible. Even if methods such as parameterisation are able to cover most future waveform modulation methods, the limitation thus lies in the front-end hardware, since true RF sampled systems cannot be expected for the foreseeable future.

In general, standards using similar modulation methods may be expected to be integrated with reasonable ease. One such example might be the integration of systems in which the baseband processing is linear, e.g., OFDM for HIPERLAN/2 and CCK for IEEE 802.11b. This is in contrast to techniques such as OFDM for IEEE 802.11a and GFSK for Bluetooth wireless technology, which are much harder to combine due to the nonlinearity of GFSK. However, due to Bluetooth wireless technology being a low cost standard, its absorption into computationally complex WLAN systems is generally not constraining, and even provides a chance for higher cost Bluetooth implementations offering considerably enhanced performance.

7 Software Aspects



By Neil Briscombe, QinetiQ.

7.1 Introduction

In this chapter we look at software aspects of software defined radio (SDR) as pursued by industry and, specifically, the Software Defined Radio Forum (SDRF). We will also consider the behaviour of networks that include elements as flexible as SDR.

There are three main thrusts to this section that cover the software communications architecture, technologies that could support realisation of cognitive networks and human machine interface (HMI) issues.

First, we cover the more conventional SDR software issues: the motivation behind the Software Communications Architecture (SCA), how SDR systems are defined through Object Management Group (OMG) models and the SCA reference implementation. During the course of our study a report was published to which we wish to draw to the reader's attention [243], where similar approaches are taken, but it focuses on different details. Here, we focus on issues that will impact commercial realisation rather than provide technical summaries.

Second, we cover software related technologies that will enable SDR platforms to extend beyond the traditional views of communication payloads towards cognitive radio and cognitive network visions. We claim that in such complex systems of systems, where a high degree of configurability and dynamism may be exhibited, system and service dependability will be both crucial and will pose difficult problems. We propose that cognitive networks will require policy based management, services could be enhanced through using the World Wide Web Consortium's (W3C's) efforts towards security and that union of wireless grids and SDR technologies would benefit each other.

The model shown in Figure 7-1 generalises these areas of research. It is presented as a simplistic model for analysing research efforts towards SDR-related cognitive networks. This is a minimum requirement for any system that could be used with any useful level of trust. Electromagnetic environment sensing is deliberately excluded as a separate entity, as this function could be distributed within networks.

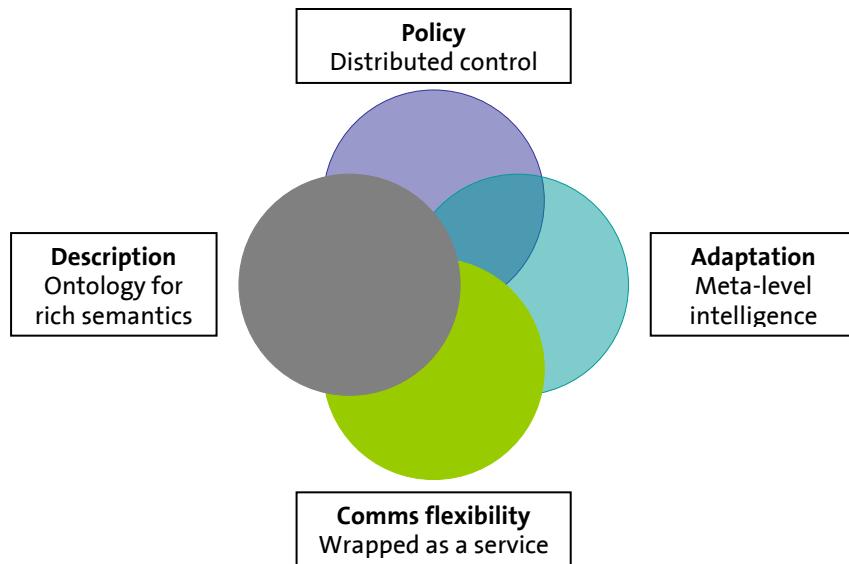


Figure 7-1: Research Model of SDR-related cognitive networks

Third, we briefly discuss HMI issues for SDR by classifying the user types of SDR and discussing the issues relating to each type.

We will conclude this section with some observations on possible areas for further work.

7.2 Open Software Architectures

In this section we will consider why it is advantageous for SDR to adopt a platform-independent software architecture, how the SCA evolved, how it affects SDR development and discuss the reference implementation by Communications Research Centre, Canada (CRC).

7.2.1 Motivation for the Software Communications Architecture

7.2.1.1 Introduction

The motivation for the SCA lies in requirements for cheaper, more effective, adaptable and reconfigurable radio communication systems to allow new developments to be efficiently exploited. We begin our discussion by considering the Joint Tactical Radio System programme. The JTRS programme was initiated by the US DoD and defence industry partners in response to shortfalls in capability and interoperability issues of deployed radio [244]. The main capabilities sought were for more agile radio platforms and the ability to exploit the increased pace of developments in existing equipment.

UNCLASSIFIED

7.2.1.2 Requirements Addressed

The JTRS programme was initiated to address the following requirements:

- Software controlled radio equipment that can be updated from remote sources:
 - Over-the-air programming
 - Downloadable waveforms
 - Reconfiguration of many communication parameters and characteristics.
- Effective and efficient to engineer – Standardised, open, modular and scaleable.
- Extend interoperability across:
 - Military, civil and commercial domains
 - Ad-hoc networks with full wireless roaming
 - Various background environments.
- Reduce total life-cycle costs for radio systems.
- Allow multiple vendor adoption, use of extant technology and new technologies as they emerge.

In an attempt to satisfy the DoD's requirements, Raytheon, BAE Systems, Rockwell-Collins and ITT formed the Modular Software-programmable Radio Consortium (MSRC).

The adoption of hardware developments in line with Moore's law, as observed by Bose [245], has already been understood to be a requirement for SDR systems. Even where systems (e.g., SPEAKeasy) have been designed to be able to be reconfigurable, they can potentially be tied to specific hardware, preventing advances in hardware from being exploited readily. Efforts must be made such that much of the software design is independent of the hardware; this is reflected by the direction taken by the MSRC.

MSRC defined the SCA, targeting the OMG standards and the common object request broker architecture (CORBA) as a means of reusing the considerable extant interoperability effort. SCA was created to define the overall architecture to allow the interconnection at the software level and real-time modification of SDRs.

The development of CORBA implementations shows that subsets of proposals are more usual as implementations. Commercial CORBA 3.1 implementations, for example, occurred years after the expected implementation timescales and the initial portable object adapters were only possible through exploitation of Enterprise Java Beans (EJB). Despite efforts for interoperability and tight specifications for communications between implementations, there can be subtle divergences that can have significant impact. Examples of this are in the level or exact nature of the implementation of defined services, differences in which can result in interoperability failure. Note that some such implementation differences are intentional and can even be driven by a commercial need to differentiate products.

UNCLASSIFIED

One could expect that SCA implementations could follow the same route, with some incompleteness and inconsistencies remaining. It should be noted that the submitted platform-independent model (PIM) does not have CORBA as a prerequisite. Despite the SCA being the de-facto architecture, it is not the only means by which the DoD requirements for SDR could have been met. However, now that it is in place, it is difficult to see how it could be replaced in the military domain.

7.2.2 Development of the Platform-Independent Model for SCA

7.2.2.1 Introduction

Here we describe the role of the OMG in SDR, namely providing a common modelling framework and reusing their significant efforts towards system integration through CORBA, in particular. The OMG's commitment to focus on producing commercial implementations of its specifications through its structure and standardisation process is also given consideration.

The OMG was established in 1989 and is the world's largest software consortium. Its mission is to help computer users solve integration problems by supplying open, vendor-neutral interoperability specifications. OMG exists to facilitate enterprise integration. The object management architecture (OMA) provides the vision and roadmap against which its other standards and architectures are authored. These standards are designed to minimise the burden of writing software bridges and generally minimise integration and maintenance issues.

Integration is addressed through a system life cycle; from business modelling to system design and component construction through to assembly, integration, deployment, management and evolution. This vision is embodied by the model driven architecture (MDA).

The scope of the organisation was broadened significantly when OMG issued several important modelling specifications including the unified modelling language (UML).

Having created the CORBA interoperability standards, OMG has, in the past, used them almost exclusively as the basis for creating standards for use in particular application domains.

The OMG board of directors votes to formally adopt specifications on behalf of OMG. Each OMG technology committee (domain and platform) and architecture board provides technical guidance to the board of directors. In addition, the business committee of the board of directors provides guidance to ensure that implementations of adopted specifications become commercially available.

7.2.2.2 OMG's Commercial Focus

OMG encourages commercial adoption of the specifications it publishes. In order for proposals to proceed through the standardisation process, they are subjected to technical as well as legal or commercial considerations to ensure that implementation is practicable. Technical reviews by the relevant OMG technology committees judge the technical merit of proposals. The OMG business committee judges legal and commercial issues and also requires evidence that a commercial implementation is either in place or likely within twelve months.

UNCLASSIFIED

The business committee has the following evaluation criteria:

- Cross platform implementation
- Commercial availability
- Intellectual property (IP) rights
- Specification publication
- Continued support.

The number of implementations relating to the CORBA specification is testament to the success of this rigours approach to commercial as well as technical issues.

7.2.2.3 OMG's Focus on SDR

Through the Telecoms Task Force, a Software Radio Special Interest Group was formed. More recently, the Software Based Communication Domain Task Force (SBC DTF) was formed. The mission of the SBC DTF is given as [246]

“The development of specifications supporting the development, deployment, operation and maintenance of software technology targeted for software defined communication devices.”

On its website, the goals of the SBC DTF are cited as [246]:

- Promotion of UML and model driven development technology in the software defined radio field
- Development of specifications to improve interoperability and exchangeability of software defined communication components
- Collaboration with other OMG taskforces on related or overlapping technology specifications
- Broaden previous Domain SIG charter with new related technologies, e.g., cognitive radio, streaming components, digital IF, spectrum management, etc.
- Promotion of OMG specifications within the software radio community
- Maintain liaison with stakeholders of software defined communication technology outside the OMG
- Provide a standard means, by which software radio-based applications are developed, deployed and managed
- Provide the capability for reconfiguration of radio networks, services, access nodes, and terminals
- Provide a standard platform-independent architecture to support software radio-based applications
- Promote the development of standard radio-based services for use by applications
- Promote the development of standard radio-based interface definitions for use in application development

UNCLASSIFIED

- Promote, as required, the infrastructure and interface definitions needed to support:
 - Transparent security solutions
 - Safety critical solutions
 - Fault tolerant solutions
 - Real-time and embedded solutions
 - Ensure compatibility and consistency of resultant specifications related to software radios.

All of the above goals have relevance to the cause of this study, but the highlights are the standards for:

- Platform-independent architecture and support for software radio-based applications
- Radio-based services for use by applications
- Radio-based interface definitions for use in application development.

7.2.2.4 Platform-Independent and Platform-Specific Models

The OMG final adopted PIM and platform-specific model (PSM) software radio components specification publication was published 1 July, 2004. Comments were due 1 February, 2005, and the recommendations were expected 22 April, 2005 (one week after the OMG meeting).

The proposal uses existing OMG models to produce both platform-independent and platform-specific implementations. These then act as references for the development of SDR as distributed software functionality. The OMG believe the proposal will prove effective at increasing the radio spectrum efficiency.

The PIM works at a layer of abstraction above middleware implementation, i.e., it does not specify the use of CORBA as a means of interfacing between the software functions of separate radio systems. The PIM allows for mappings to be made to other platforms that could use lighter-weight middleware solutions, for example. Tools can be used to automate test interoperability. The OMG MDA guide uses the example of a CORBA EJB mapping translation, but the premise applies for any PSM drawn from the same PIM.

There is academic work [247] to support the essentially commercially driven approaches of the MDA modelling for SDR and the benefit of the mapping of PSMs to the SDR PIM.

If widespread commercial or even domestic adoption of SDR is realised, it is likely to be through cheaper and more modestly specified equipment than that found in the military domain. However, as the PIM is technology neutral, PSMs may be adopted for the more humble equipment that will inevitably find its way into our everyday lives. The deciding point for this is likely to be the cost/benefit analysis for exerting extra effort on interoperability of wireless devices against the perceived priorities that users have.

UNCLASSIFIED

PIM/PSM mappings are one identifiable way in which the SDR community has attempted to meet the DoD requirements for interoperability and reduce development time and costs. Real benefit can only be drawn from these if there is an open, well-understood reference implementation that can stand the rigors of the mapping and to blaze a trail and act as a case study.

7.2.3 Reference Implementation of SCA

7.2.3.1 Introduction

CRC is Canada's primary research and development facility for advanced communications. Two of its main concerns are terrestrial wireless and satellite communications. Its activities support native defence policy and the national economy through technology transfer.

In partnership with the Defence Research and Development Canada (DRDC), and with support from the SDRF, CRC created what has become the de-facto reference implementation.

CRC were involved in the evolution and adaptation of SCA specifications and remains instrumental to the development and commercial uptake of the SDRF. CRC cite their SDR contributions as [248]:

- Pioneered the use of a dynamic loader for DSP environments and the use of SCA CORBA adapters
- Developed and tested a SCA-enabled radio demonstrator, supporting a DSP-based implementation
- Released on its web site, an open source reference implementation of the SCA (project (SCA reference implementation) SCARI-OPEN sponsored by the SDRF)
- Submitted over 20 change proposals to the JTRS/Joint Program Office to enhance the SCA specifications
- Carried out the first public demonstration of a commercial SCA-enabled waveform, a digital audio broadcast (DAB) waveform.

The SCA specification required a codified implementation to bring credence to the standards. Experience of incompatibility issues with CORBA object request brokers showed that an open source reference implementation could serve as a tangible case study as well as providing auditable conformance tracing for other PSMs. Also, the process of engineering should, and did, lead to useful changes to the standards in regards to their practicability.

Since the arrival of the reference implementation, SDR vendors now have the opportunity to reuse CRC's SCA implementation, reducing the overall effort of developing SDRs and improving their chances of interoperability (assuming that the reference implementation is widely adopted). The SDRF's support will be instrumental in any adoption of the reference model and the reference model is intended to be instrumental to the commercialisation of the technology.

UNCLASSIFIED

CRC chose to first implement the SCA reference implementation (SCARI) in the Java programming language. The main reason for selecting Java was its platform independence, making it easy to deploy on a development platform of the vendor's choice with no issues relating to multiple platform releases and configuration of make files, for example. Other benefits included its original design goal of object orientated embedded control, giving it all necessary designing concerns and syntax and also the removal of low-level abstractions such as memory management and pointer math.

7.2.3.2 Additional Features

The entire SCA core framework is covered in the CRC reference implementation. Desirable features other than the core framework that are also codified include the extensible mark-up language (XML) domain profiles, service interfaces and sample waveform applications.

Also offered by CRC are tools to support the operation of other SDR implementations. Points of interest to this study are the inclusion of a node manager and a 'waveform application factory'. The node manager initiates nodes and their constituent components. This includes support for logical devices, services and supports service discovery through look up in the naming service. The application factory selects valid implementations for each component of the SCA taking into account interdependencies between different components. Again service support is provided as is discovery through the naming service.

CRC provides UML documented diagrams, code examples and a general purpose inspector, which works for any SCA component through the Java reflection/introspection mechanism.

7.3 Technologies towards Cognitive Networks

SDR brings the vision of cognitive networks of SDR nodes sharing resources such as spectrum. In this section we look at the issues associated with this vision from a software point of view and then suggest some existing research areas that could combat these. We start with broad notions of dependability, focusing on how these can be managed by policies, look to W3C standards as a means of providing a secure service-orientated architecture and scope the development of wireless grids that we assume will be the first realisation of cognitive networks.

7.3.1 Managing Faults in SDR Networks

Here we explore the management of faults in networks with SDR nodes. We attempt to address issues beyond those related to non SDR networks [249]; communications payloads that are re-configurable over the air produce unique problems.

UNCLASSIFIED

Dependability is a fundamental problem for SDR and cognitive networks. Despite efforts to make SDR systems secure, and in keeping with the reality of network security practice, it should be assumed that SDR devices *will* be vulnerable to attack (note that worms for attacking mobile phones have already been written and released ‘into the wild’). The aim of any deployed security mechanisms is to limit damage and to maximise whatever assurance might be practically achievable. Related to the assumption that security vulnerabilities will exist is the assumption that faults will occur. If a radio downloads and installs some software, it will typically be very hard to *prove* that the software is reliable. As such, it must be assumed that the behaviour or interaction of the new software will, in certain instances, cause faults. More generally, given that radios are interacting with an unknown environment, the occurrence of failures, due to a range of possible reasons, must be taken into account. In this respect, an SDR-based cognitive network will have much in common with any networked information infrastructure, where any of a large number of reasons might cause downtime.

Before discussing what the dependability challenges might be, some clarification of terminology is useful. In the dependability literature, distinctions are made between *fault*, *error* and *failure*. A *fault* is defined to be the underlying (and often latent) cause of a failure. An example of a fault might be a bug in some software. When the fault is invoked and manifests itself, e.g., the bug is executed, then that is an *error*. The consequence of the error is the *failure*. Continuing our example, if the software with the bug is controlling a server, then the consequence of the error, *viz*, the failure, might be the server crashing. It is important to realise that the distinction between failure and fault is relative, which is to say that whatever constitutes a failure at some given level of abstraction can be regarded as a fault at another level. Continuing our example, if the server is a member of a server farm, the failure at the level of a given server then becomes a fault at the level of the server farm.

In our work we depart from the above, time-honoured, notions. Our perspective is rooted in the end-user as opposed to the technology-centric perspective of the fault, error and failure definitions. Thus we only consider failures, which might have many diverse causes, but have the underlying notion in common that a failure marks *the legitimate expectation of an end-user not being met*.

In an SDR context, there are many possible types, or instances, of failure that are in fact enabled by the very flexibility of SDR and cognitive radios and networks. Examples might include:

1. A worm penetrating firewalls and infecting devices. In conventional networks, the threats posed by worms has only grown and, given that worms already exist for mobile phones, it is unlikely that they will not be a threat in cognitive SDR networks. The damage that might result from a worm is very significant, what that damage will be depends on the worm’s payload.
2. Problems caused by downloading flawed software. The failures that users experience on a day-to-day basis are likely to have benign and mundane causes. If software is downloaded to upgrade a device, that software might fail to install for any of a number of reasons. The result of this might be either the temporary loss of service or, possibly, the device crashing.

UNCLASSIFIED

3. Failure to achieve agreement on spectrum allocation. This could be for many reasons, both benign and malicious. If one of a set of radios legitimately upgrades itself with benign code, the consequences of that change of functionality might be such that spectrum agreements fail, or existing agreements are violated. A more dangerous possibility is that a malicious device manipulates agreement protocols for its advantage, or to inflict denial-of-service.

With the above assumption, definitions and the need to achieve dependability in mind, the problems to be addressed, principally through the use of policies, are:

- How to minimise the occurrence of failures
- How to facilitate (and maximise the speed of) recovery from failures, given that they are bound to occur.

Note that these ultimately relate to minimising the total ownership cost of the technology.

As we will explain later, a significant problem in the use of policies is that of finding the best policy. For dependability, the strategy that we recommend for developing policies is to adapt and abstract mechanisms, such as fault tolerance protocols, from the traditional dependability community to work at higher levels of abstraction. Instead of a protocol that defines each action that must be taken, we use the same protocol as a basis for defining a policy, probably an action policy, which will guide agents in the way in which they respond to failures.

7.3.1.1 Minimising the Occurrence of Failures

The occurrence of failures can essentially be prevented in two ways:

1. To eliminate the root cause of the failure. The means for achieving this will, of course, be determined by the underlying cause of failure. For instance, in the case of failures caused by bugs in code, either the code quality has to be improved so the code is ‘bug-free’ before deployment or bugs have to be identified and ‘patched’ before code deployment. In the case of security breaches or failures, the failure elimination techniques are the security mechanisms and policies used.
2. To mask a failure when it does occur. Failure masking requires continual, proactive mechanisms to check whether an error has occurred and then to ensure that the failure does not manifest itself. From traditional dependability, the concept of fault-masking is highly germane to our interests here. Fault-masking is achieved by retaining functionally active replicas of a component and then enforcing a checking mechanism across the replicas. This ensures that errors are identified and prevented from being manifest.

UNCLASSIFIED

7.3.1.2 Making Recovery as Fast and as Easy as Possible

Our basic philosophy is to assume that, despite best attempts at minimising their occurrence, failures will occur. Thus, the point is how do we ensure that recovery is as fast and as ‘painless’ as possible? In traditional dependability, recovery is achieved by various means of fault detection, supported by a means to return to a normal state. The return to a normal state is typically supported by techniques such as checkpointing, in which the system’s state would be stored periodically as a record from which to return to normality. These techniques are generally referred to as ‘rollback’ recovery.

In recent years these same ideas have been taken forward in an innovative way by the recovery oriented computing (ROC) work at UC Berkeley [250]. In effect, this work lifts rollback recovery mechanisms to a semantic, policy level. The essential idea is to implement operators that allow a system to *rewind*, *repair* and *replay*. In other words, when a failure is detected, *rewind* to the point of origin of the failure, *repair* the failure and then *replay*, by restarting the system. The aim of the ROC approach is to make recovery as fast and as easy as possible. A similar approach would be of benefit to cognitive SDR systems.

7.3.2 Controlling SDR Interactions through Policy Based Control

7.3.2.1 Management and Control Problems in SDR

The overall aim of a cognitive SDR system is to achieve a reliable and efficient communications service. The alleged advantages of cognitive SDR are that radios are far more adaptive and responsive, principally in the following regards:

- The functionality they offer and implement. For example SDR devices can download software upgrades on-the-fly
- The way cognitive SDR devices interact with their environment. This is usually understood to refer to spectrum allocation, but it can be far more complex than that. Potentially, radios can modify the direction in which they transmit, radios can tune MAC layer agreements and, in an ad-hoc or delay/disruption tolerant networking paradigm, SDR-enabled devices might perform routing services. All these various aspects of resource management are necessary to support the appropriate quality of service required by the application.

Furthermore, it should be noted that these adaptive radios are largely autonomous. Typically, the SDR devices will not even belong to the same domain. The combination of flexibility and adaptiveness, as indicated above, along with autonomy of control then creates significant challenges for co-ordinating resources (of the various types mentioned) and ensuring that the communication service offers a basic reliability. We propose to address these challenges using *policies*.

UNCLASSIFIED

7.3.2.2 What are Policies?

Policies are mechanisms for controlling the behaviour and interaction of autonomous (or partially autonomous) entities, henceforth referred to as *agents*. Note that a cognitive, SDR-enabled device is, in effect, an agent by virtue of its capacity for independent behaviour and self-adaptation. In essence, policies allow agents to be controlled without a full prescription of behaviour for all scenarios that the agent might encounter and without requiring continuous monitoring of all that an agent does. As such, policies are fundamental to achieving and controlling autonomy. The growth in complexity of systems is driving the current interest in autonomy. This in turn is motivating the interest and application of policy-based control. There are many applications of policies, including QoS, access control, business processes, communication and many more. Of these applications, an interesting class are policies to manage trust or achieve some notion of dependability in complex systems. A good example of policies to achieve dependability on complex systems is road traffic rules, which are, in effect, policies designed to minimise accidents and expedite traffic flow.

Policies themselves can be broadly characterised into the following types:

- Action policies: These policies guide an agent in performing some sequence of events. In performing the indicated events, the aim is to transform the state of the system (comprising of a set of agents and their environment) from some given state to some desired state. Note that, however, action policies are relatively prescriptive. They will partly specify the actions required to reach a state. Action policies are usually guarded by some monitor on environment (and agent) state.
- Goal policies: Action policies include some prescription of events that have to be undertaken to achieve some objective. Goal policies simply specify that objective and let the agent decide how it might best be achieved. In order to use goal policies, agents must be endowed with the wherewithal to derive the actions required to achieve the goal.
- Utility policies: In goal policies, the goals will themselves be defined with some high-level gain in mind. Thus it is possible to specify policies at an even higher level than goals by simply defining the required gain. This is a utility policy.

Before discussing what we might want to achieve with policies, it is important to be clear about the challenges associated with their use.

7.3.2.3 Key Problems in the Use of Policies

In using policies there are certain basic challenges which are generic to any use of policy. In the domain of SDR and cognitive radio, these same challenges manifest themselves in their own specific way. The generic challenges are:

1. What is the right or ‘best’ policy? This is a fundamental question in the use of policy. Related key questions include how the best policy is found and how policies are evaluated to decide which is best. Whilst there is significant discussion about the use of policies, the problem of finding the right policy, which is arguably the crucial question, receives far less attention.

UNCLASSIFIED

2. How is policy enforced or implemented? This problem assumes that the right policy has been found. In such a case the policy may constrain the activity (or resource allocation, etc.) of agents. If so, it may be in the self interest of some agents to violate or circumvent a given policy rather than adhere to the policy's intent. Thus, the problem of enforcing policies is clearly important. Alternatively, an agent will often need to determine a course of action in order to meet or follow a given policy. In this case the agent will need to be able to evaluate possible options. This might involve the agent solving an optimisation problem, which can be computationally expensive.
3. How is policy distributed? For a policy to be implemented it has to reach all the agents that need to be aware of the policy and that are responsible for implementing it.
4. How are policy conflicts managed? If agents belong to different domains, it is unlikely that the different domains will have identical policies. When the agents interact, it will be necessary to resolve policy conflicts.

In the case of SDR and the cognitive radio/network vision, the above challenges are compounded by the following constraints:

1. Multiple domains with no centralised control. SDR devices cannot be assumed to belong to a single domain. Therefore, the policies which they attempt to respect are not likely to be the same. Thus the problems of policy enforcement and policy resolution must be solved without a single controlling infrastructure.
2. Uncertainty. Radios operate and configure themselves in an environment of significant uncertainty. The elements of this uncertainty include which radios exist in a given location, their configuration and so forth. Further sources of uncertainty include the fact that the radios in a given environment cannot be assumed to be trustworthy. Nor can the trustworthiness of software that is downloaded for reconfiguring a device be guaranteed. With regard to the policy problems mentioned above, the uncertainty will make it difficult to decide what the best policy is.

In summary, while the prospects promised by the flexibility and adaptiveness of cognitive SDR are very alluring, realising them requires solving significant problems concerning the coordination and control of distributed, autonomous, self-adapting agents. A fundamental element in this coordination and control are policies.

7.3.3 SDR Applicable Techniques from the World Wide Web Consortium (W3C)

7.3.3.1 Introduction

The W3C was founded by Tim Berners-Lee in 1994. The W3C now has over 350 members and is responsible for the open protocols and interoperability standards that underpin the World Wide Web. It cites three principal tasks [251]:

1. Vision: W3C promotes and develops its vision of the future of the World Wide Web. Contributions from several hundred dedicated researchers and engineers working for member organisations, the W3C team (led by Tim Berners-Lee, the Web's inventor) and from the entire Web community enable W3C to identify the technical requirements that must be satisfied if the Web is to be a truly universal information space.

UNCLASSIFIED

2. Design: W3C designs Web technologies to realise this vision, taking into account existing technologies as well as those of the future.
3. Standardisation: W3C contributes to efforts to standardise Web technologies by producing specifications (called ‘recommendations’) that describe the building blocks of the Web. W3C makes these recommendations (and other technical reports) freely available to all.

The commonality between the OMG and W3C is system interoperability. The OMG is focused on ensuring the supply of implementations over multiple platforms through the robust mapping between models. The W3C is focused on Web based delivery. This divergence in focus is reflected in the standards that emerge from the two organisations. The emphasis for OMG ultimately results in multiple PSMs with automated or assisted mappings between them. For the W3C, the result is protocols to enable Web based connectivity.

Initial use of the W3C centred on the common presentation of electronically published material. The W3C team had grander visions for the World Wide Web with the notion of the Semantic Web. W3C defines Semantic Web in these terms [252]:

“The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the resource description framework (RDF), which integrates a variety of applications using XML for syntax and URIs for naming.”

The above statement is made in terms of W3C standards, but it can be thought of as a general vision for the sharing and interpretation of information between machines. With regards to the Semantic Web, the literature often focuses on the use of ‘agents’ in this process. Agents are usually considered as intelligent, adaptive, often mobile, software entities working asynchronously on behalf of humans. Though such complex agents may well have an important future role in SDR networks, other forms of software such as simpler control systems can play their part in the formation, use and management of SDR networks in the near term while still working towards the Semantic Web vision. Below is a higher level statement from Berners-Lee et al [253]:

“The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.”

Increasingly, this vision is being realised as more Web-based inter-application interaction occurs. Web services refer to the interfaces made available to enable this interaction. W3C defines a web service as [254]:

“A Web service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Other systems interact with the Web service in a manner prescribed by its description using SOAP messages, typically conveyed using HTTP with an XML serialisation in conjunction with other Web-related standards.”

UNCLASSIFIED

7.3.3.2 Web Service Architecture

With a broad interoperating remit covering many types of system, W3C defines web services in the web service architecture (WSA). This is an example of the computing service orientated architecture (SOA) paradigm which is attracting increasing interest and can easily map on to the efforts of the OMG. As such, the WSA works at a high enough level of abstraction to be useful to this study. The WSA does not attempt to specify implementation details and covers only the minimal and common characteristics.

The WSA has four models contained in an overarching meta-model. These models are:

- Message oriented model (MOM)
- Service oriented model (SOM)
- Resource oriented model (ROM)
- Policy model (PM).

These are discussed separately below.

7.3.3.3 Message Oriented Model

The MOM focuses on the structure of messages, how they are transported and message processing, but not their content. The model consists of thirteen developments, but is a useful reference mechanism for considering communications between applications. It is simplistic enough to be used by those not familiar with communication system engineering, yet it contains some powerful concepts such as message exchange patterns, message reliability and delivery policies.

The concepts covered in the MOM are a fundamental part of the WSA. They can be readily coded to WS specifications, are supported by WS centric software engineering tools and are extremely useful to engineering SDR systems and applications hosted therein. For example, a WS software engineer would be able to interface with these entities and specify the way in which security or QoS could be negotiated between applications. This could be implemented very simply by using them as configuration datum for control by users or administrators. The MOM and its supporting specifications and tools will provide the framework within which software engineers might work towards specific implementations. There is the potential for automatically mapping through the MOM to specific implementations of its standards to further simplify the construction of interoperable systems towards the Semantic Web.

7.3.3.4 Service Oriented Model

The SOM focuses the next higher level of abstraction from the message: a requested for action in the real world. In the WSA the ‘real world’ is modelled as a person or organisation (owner – those responsible for the service), service task, service role, and goal state entities.

UNCLASSIFIED

The SOM makes use of metadata, which, as described in the service oriented architecture, is a key property of service oriented architectures. These metadata are used to document many aspects of services, from the details of the interface and transport binding, to the semantics of the service and what policy restrictions there might be on the service. Providing rich descriptions is essential to the successful deployment and use of services across the Internet.

The MOM is useful to engineers of SOAs in that it allows one to concentrate on service request issues rather than on the underlying messaging that is required. The SOM can allow remote systems to be viewed as if they are local.

7.3.3.5 Resource Oriented Model

The ROM focuses on those aspects of the architecture that relate to resources. Resources are a fundamental concept that underpins much of the Web and web services. A web service is a particular kind of resource that is important to this architecture. The resource model is adopted from the Web architecture concept of resource.

The ROM focuses on the key features of resources (independent of the resource's role in web services). Issues such as the ownership of resources, policies associated with resources and so on are covered.

7.3.3.6 Policy Model

The PM focuses on constraints on the behaviour of agents and services. We generalise this to resources, since policies can apply equally to documents (such as descriptions of services) as well as active computational resources.

Policies are 'about' (i.e., refer to) resources. Policies are applied to agents that attempt to access shared resources. Policies are acted on by software controlling resources, to ensure proper services levels and actions in the real world. Policies can have aspects that relate to security, QoS, general management/configuration, network management or application specifics.

The potential importance of the Semantic Web for SDR centres on the cooperation between SDR transceivers and also between the users of applications that communicate across them. Developments made for the interoperability of systems should be relevant provided that they are engineered at high enough levels of abstraction.

Before introducing individual standards, it should be noted that a cost benefit analysis is required to ascertain the possible impact for the users. Assuming the applicability of the use cases for SOAs to wireless networks, this cannot be addressed by the use of traditional QoS metrics such as bit rates and latency issues. Instead, analysis needs to be centred on the user's requirements and perceptions of the services delivered. There will be traffic management overheads associated with a SOA and its resulting specifications and protocols. However, their role in negotiating a secure network for end-to-end services, for example, could be a necessary precursor for ad-hoc networks to form and provide anything useful to their users.

UNCLASSIFIED

While there is work on the traditional notions of QoS for SOAs [255], work on the emerging standards landscape with a higher level of QoS abstraction has yet to be completed. However there is a study [256] on the use of compression schemes for particular application to XML in wireless environments that greatly improves matters for the possible application of such standards.

7.3.3.7 The WS Standard Family

As will be argued in Chapter 8, issues relating to security and the user's confidence in it will be a significant set of problems that will need to be addressed in order to assure the long-term adoption of SDR. The W3C is coordinating the efforts of the world's largest software suppliers towards a set of open and comprehensive standards. These tackle the difficult challenges that are faced when forming and managing SOAs, allowing the negotiation of security issues and the resolution of conflicts that arise in ad-hoc arrangements.

7.3.3.8 Summary

The WS-* specifications as a whole provide building blocks for web and grid services, allowing interoperability and platform/language independence. With the distributed nature of web and grid services, security considerations are of paramount importance and the problems raised are not easily solvable with current technologies. The security-related WS-* specifications, described in Appendix C, have been proposed to address the potential problems that will be faced by developers and users of widely distributed web and grid services, and to address issues of trust and reliability.

The standards placed higher in the family's model provide the most potential from a functionality viewpoint. However, there remain challenges if these are to be successfully implemented and adopted. Despite these challenges, the SDR community would benefit from considering the approaches of WS-Security and WS-Policy in SDR scenarios, as they deal with trust ambiguities associated with ad-hoc networks.

Rather than being yet another burden on SDR implantations, the SOA, the WS-Security family and implementations therein could prove to have mass produced implementations sooner, and on 'thinner' platforms, than through the SCA. The truth of this premise depends largely on the definition of SDR (i.e., if one can accept SDR if not being an implementation of the SCA). It might be the case that a significant portion of the submitted PIM for SDR could be realised by means such as those being developed for web services.

7.3.4 SDR and Grid Computing

7.3.4.1 Introduction

Grid computing heritage draws from parallel and distributed computing. Its emphasis is on the controlled sharing of resources. The resources in question were traditionally computational. This provided the ability to use remote processing power. Grid taxonomies [257], however, are becoming broader. Grids for (often academic) computational resource sharing are referred to as 'file/compute grids' but also include, amongst others, information, complex, enterprise, campus and autonomic grids.

UNCLASSIFIED

Before considering the broader aspects of resource sharing it is worth noting that if SDR finds more commercial realisation, it is likely to find exploitation routes that increasingly include less well as well as better specified equipment as it gains acceptance. If this is the case, then the file/compute grid type may prove to be a useful model.

As noted above, grid research increasingly encapsulates broader issues than that covered by file/compute grids. The motivation to enabling e-business issues has brought importance to the notion of the virtual organisation (VO). A VO is an arrangement composed from existing, formalised organisations to better achieve a common goal. The issues surrounding the formation and management of VOs has become the main research challenge for grid research.

7.3.4.2 Grid Computing

This boarder notion of grids is also suggested in earlier work that focuses on computer resource sharing [258]. Breadth is derived by means of the full list of resource sharing cited as being aimed for. In a non-exhaustive list, network resources, code repositories and catalogues were also included, showing the grid community's desire to share network (and communication) resources.

In relation to network resources management, mechanisms that provide control over the resources allocated to network transfers, their prioritisation, reservation and enquiry functions to determine network characteristics and load are understood as necessary. The management of networked resources is considered in the collective layer.

The list below, cited by Foster [258], expressed as resource sharing for VOs, is equally important to wireless grids:

- Directory services allow VO participants to discover the existence and/or properties of VO resources. A directory service may allow its users to query for resources by name and/or by attributes such as type, availability, or load.
- Co-allocation, scheduling, and brokering services allow VO participants to request the allocation of one or more resources for a specific purpose and the scheduling of tasks on the appropriate resources.
- Monitoring and diagnostics services support the monitoring of VO resources for failure, adversarial attack ('intrusion detection'), overload and so forth.
- Data replication services support the management of VO storage (and perhaps also network and computing) resources to maximise data access performance with respect to metrics such as response time, reliability and cost.
- Grid-enabled programming systems enable familiar programming models to be used in grid environments, using various grid services to address resource discovery, security, resource allocation, and other concerns.
- Workload management systems and collaboration frameworks – also known as problem solving environments (PSEs) – provide for the description, use, and management of multi-step, asynchronous, multi-component workflows.

UNCLASSIFIED

- Software discovery services discover and select the best software implementation and execution platform based on the parameters of the problem being solved.
- Community authorisation servers enforce community policies governing resource access, generating capabilities that community members can use to access community resources. These servers provide a global policy enforcement service by building on resource information and resource management protocols (in the resource layer) and security protocols in the connectivity layer.
- Community accounting and payment services gather resource usage information for the purpose of accounting, payment and/or limiting of resource usage by community members.
- Collaboratory services support the coordinated exchange of information within potentially large user communities, whether synchronously or asynchronously.

7.3.4.3 Wireless Grids

The idea of wireless grids comes from the culmination of three computing paradigms, namely, grid computing, peer-to-peer (P2P) computing and web services, whereby services can be discovered and called upon in a mixed wired and wireless network. Networks for wireless grids are typically considered to be ad-hoc, which raises challenges relating to trust and dependability of service provision. The inclusion of P2P technologies allows operation with lower trust requirements and to deal with service disconnects, etc.

The following is a simple (and not mutually exclusive) classification for wireless grid applications [259]:

- Applications aggregating information from the range of input/output interfaces found in nomadic devices.
- Applications leveraging the locations and contexts in which the devices exist.
- Applications leveraging the mesh network capabilities of groups of nomadic devices.

This work points to the need for understanding issues for shareable resources, where they are used, who owns them, how they are controlled and how they are monitored remotely. There are several precursors to the provision of resource sharing, these being resource description, resource discovery, coordination, trust establishment and clearing.

Resource description and discovery are problems common to all SOAs. Efforts in the web services community such as UDDI and OWL-S facilitate online databases of available services.

Coordination systems allow the controlled (often scheduled) sharing of resources. Mechanisms are often tailored to resource types such as disk space or available clock cycles. A more general approach needs to be adopted for SDR. Here we can look to the web services community, which will need to address the same issues when implementing resource sharing under OWL-S.

UNCLASSIFIED

Portable and nomadic SDR devices will bring their own challenges that need to be addressed. In addition to the general ad-hoc networking issues, the reduced capabilities due to size and weight constraints imposed will result in far thinner platforms.

Thinner nomadic SDR platforms would be able to provide enhanced service provision if they could discover and scavenge resources available over the air. This points to the importance of wireless grids to SDR as a whole. The main challenges that remain are the establishment of trust and its negotiation in times of conflicts in requirements and circumstance.

Proposed models for positioning technologies related to this area [259] show cause for concern in the amount of overlap and competition that exists. Many challenges remain if ad-hoc wireless grids are to be formed and deliver any useful services. The remaining challenges include harmonisation, bridging and federation such that networks can be managed, understood and trusted by their users.

The convergence of web services and grid services (managed by the Global Grid Forum) should improve matters somewhat. When this has been resolved, the resultant body of standards should make a useful starting point for grid-over-SDR research.

7.4 SDR and Human Interface

7.4.1 Introduction

In this section we will deal with human machine interface issues for SDR. We will classify the SDR user types and discuss the issues relating to each type briefly.

7.4.2 Identified Users

We begin by identifying various types of user. For the purpose of this study, three types of user have been identified. These are:

- Those people using SDR equipment in the field.
- Those people responsible for forming or managing networks of SDR equipment.
- Those people who create and maintain the software elements in SDR systems.

7.4.3 HMI for SDR Equipment

The HMI will need to operate at several layers. Primarily, HMI will be application specific, i.e., it will relate to the main reason for requiring a communications link. For example, for a video link, the interface will relate to the display of video. If SDR can bare the promise of open systems that are less handset vendor or network specific, it is easy to understand the utility in having a common HMI for some of the layers underneath.

UNCLASSIFIED

HMI for handset configuration should be considered that allows modification of the concepts identified as enablers for SDR.

7.4.4 HMI for SDR Networks

Two obvious levels of interface can have application at this level. These are distributed monitoring and control of the networks, where some form of awareness or administration is required. Even where control is affected by the collective (e.g., using P2P approaches) there is a need for network monitoring, for example, so that specific metrics such as availability can be factored and used to improve overall service provision. Network awareness portals could be a solution for such situations where the increased traffic overheads are acceptable. In networks close to saturation this may not be the case. Here, users will have to rely on local inference and exploit information available from received packets.

Trust within SDR networks will have aspects of human centric trust as well as technical issues. Information relating to social issues may be difficult to encode and manage, though some work such as Semantic Web trust frameworks [260] could be leveraged. This would include the intentions of the users of nodes, their previous behaviour and so forth. When applied to networks containing SDR elements, trust impactions of distributed network management functions will add complexities that may not be possible to address with existing frameworks and may best achieved through human judgment and decision aides rather than autonomy. Related to this is the creation and management of policy. HMI for policy management, as part of distributed control, will be vital for the overall confidence in deployed systems.

7.4.5 HMI for SDR Development

Toolsets based on OMG standards such as UML are an ideal platform for developing SDR applications. There are many implementations based on the set of models in OMG UML. These include tools that can generate interface definition language (IDL) from graphic representations and skeletal code from the IDL. It is at this point where specific implementations and their application programming interfaces (APIs) need to be understood by practitioners with a sound understanding of both software engineering and digital telecommunication systems.

CRC's SCARI-Open includes:

- Graphical HMI tools to simplify the installation and control of SCA applications
- Basic component inspector graphical user interface (GUI) for introspections of all SCA components
- Basic application manager GUI to install, load and control waveform applications.

These toolsets provide a useful case study to judge suitability for presentation and the practical application of tools to help simplify and automate SDR development and services. The tools attempt to reduce the level of expert knowledge required from both software and digital telecommunications engineering.

UNCLASSIFIED

There is still much work to be done before SDR capabilities become a wrapped local service and made exploitable at runtime through ontologies, introspection and Semantic Web-like interactions. If this can be achieved, however, some of the major challenges (such as human centric trust) can then be directly mapped on to, and tackled by, higher level applications.

7.5

Conclusions

In this chapter we have shown that the current approach for the development of platform-independent models, based on the OMG models, provides useful levels of abstraction for software engineers as well as communications specialists. Furthermore, the W3C web services architecture provides detailed models that cover aspects important to SDR. These are the message orientated model, the service orientated model, the resource oriented model and the policy model. These can be exploited in the commercial development of services and applications that utilise SDR, including automatic PSM mapping (if adequate tooling and third party APIs are available).

Another means by which these models might be used is at runtime to allow flexible configuration, higher levels of HMI and even machine cognition based on introspection of metadata, especially where these are supported by specialist ontologies. The future consequence of this, if current research threads are successful, is that they should result in systems that are more capable of expressing what optima could be strove for, the means by which to aim for such goals, how to cooperate with other equipment, and improve HMI so they that are sensitive to new contexts encountered at runtime.

There are inherent dangers in such flexibility if a means can be found to exploit them. This is one of the drivers for controlling possible adverse behaviour emerging in SDR networks through the use of policies. Policy based networks can exert control over their elements to improve security, whilst still allowing the principals of cognitive radio to be realised [246]. However, even if control and user data are transferred using different waveforms, the possibility remains that if a security fault occurs there is significant potential for this to propagate through the network without careful policy management.

Recently there has been some advancement towards how Semantic Web technologies [261] can improve the understanding of flexible transceivers. This work is still focused on cognition relatively low in the ISO-OSI stack. However, the principals could be reapplied at a higher level, e.g., for resource sharing, service level issues or the management of policy. Also, the work is using rather simplistic configuration concerns and is reactive rather than adaptive or proactive (it is akin to simple action set policies rather than goal based policies).

The shared resource identified in grid computing, when applied to an environment of flexible wireless transceivers, opens up a great potential to enable networks of nodes that are able to share and better manage their capabilities for the communities they serve. The remaining problems of self interest and mixed trust can be tackled with multiple techniques, including (but not limited to) semantically rich P2P resource trading and policy based control. All these are current areas of research. However, the more likely risks to the promise of cognitive SDR grids are commercial and legislative rather than technical.

UNCLASSIFIED

There are also some frameworks defined for the commercial application of wireless grids [262]. However, compelling, convincing examples have yet to emerge.

If realised, wireless grids and networks containing SDR elements are likely to increase consumer call for ever more capable, flexible, adaptive, interoperable and service-aware devices that can deliver capabilities and infrastructure as they are taken up. These then are potentially available for commercialisation in a model similar to that seen in the development of the World Wide Web. Standards bodies such as OMG, W3C, and OASIS will be central to coordinating and focusing effort towards such ends.

Potential benefits of ad-hoc networks containing SDR elements to the UK economy and its e business aspirations include:

- New global markets in supplying equipment and development tools that UK companies can sell in to
- New markets in supplying services over wireless infrastructures, e.g., location based services
- More efficient and practical application of virtual organisations, especially those involving mobile staff or engaged in e-business
- More efficient use of the available RF spectrum.

8 Security



By Daniel Bradford, QinetiQ.

8.1 Introduction

In this chapter we look at the possible security requirements for software-defined radio (SDR) system. First the fundamental aspects of security are considered by defining four general requirements for security. These are analysed and appropriate technologies are identified that can be used to meet these requirements.

A specific threat analysis is also detailed which identifies possible threats to SDR. This analysis includes two dimensions which are the possible classes of perpetrators and the likely types of security violations.

The types of traffic that will be seen in SDR are divided into three categories, each with potentially different security requirements. These categories are then analysed and appropriate technologies are identified that could be used to provide the required level of security. Throughout this section, parallels are drawn to similar technologies such as GSM, public switched telephone network (PSTN) and conventional Internet protocol (IP) networking.

Finally in this section, we look at why the architecture of a general purpose processor is fundamentally insecure and hence why it might not be suitable for a system in which a high level of security is required. The issues revolve around the insecurity of the memory structure. Several technologies are identified that might help to make the memory structure more secure.

8.2 What is Security?

We begin by defining what we mean by security. The security of a system is the extent of protection against some unwanted occurrence, such as the invasion of privacy, theft, the corruption of information or physical damage. It is often useful to divide security into different system requirements. Typically this is done by dividing the requirements into four areas, namely, authentication, integrity, confidentiality and availability. We will now consider each of these in turn.

8.2.1 Authentication

Authentication is the process of verifying an identity claimed by, or for, a system entity [263].

The requirement for authentication implies that a user or device at one end of a communications link does not trust that the party on the other end of the link is who, or what, they claim to be. There may be a requirement for one way authentication, where it is only necessary to authenticate one end of a link (e.g., a web server), or mutual authentication, where neither end is trustful of the other (e.g., a virtual private network (VPN) connection over the Internet).

UNCLASSIFIED

As well as authenticating the relevant parties at the start of a session, it may also be necessary to authenticate *all* of the data transferred throughout the duration of the session. This is to stop an attacker from injecting malicious data during a session after initial authentication has been successfully completed.

A key technology that has enabled users to authenticate themselves and the data they transmit is the cryptographic hash. Common examples of hashing functions are message digest algorithm #5 (MD5) [264] and secure hash algorithm 1 (SHA1) [265]. A hashing function is a one way function which creates a new output every time the input changes. The fact that it is a one way function means that it is impossible to determine the input simply by looking at the output. Note that this is different to normal encryption which is designed to be reversible with knowledge of an appropriate key. The input to the hashing algorithm will be the data that will be transmitted and a secret key known only to the two parties involved (the secret key is never transmitted in plain text). When data are received, the user takes the input, adds the secret key and performs an identical hash. If the result is identical to the received hash the receiver can be confident that the data has not been altered in transit. As a bi-product, it also adds some integrity; if the data have been corrupted, the hash will not be identical and the test will fail.

Often this hash is encrypted with the transmitter's private key creating a 'digital signature'. This hash can only be unencrypted with the public key of the transmitter. This proves that the transmitter knows the private key associated with an advertised public key.

There is one type of attack which poses a significant problem with establishing authenticity of data known as a 'replay attack'. This is where an attacker records a message between two devices and then replays it at a later time to cause some sort of ill effect. The message might have measures in place that enable the recipient to confirm that it was originally constructed by the genuine sender. However, as it was originally constructed legitimately, it will pass the authentication tests. To overcome this problem, some sort of sequence number is added which restricts how long the current packet will be valid for. This is what is done in IP security (IPSec), a set of protocols developed by the IETF to support secure exchange of packets at the IP layer, but the replay protection is still regarded as weak.

8.2.2 Integrity

Integrity is the knowledge that data has not been altered since it was originally produced. Integrity is a potential requirement for many different reasons. For example, if a software update was being transferred, corruption of data could have many implications, ranging from a system crash to operation outside normal operating parameters. The data may have also been altered maliciously by someone attempting to affect the system in some way.

UNCLASSIFIED

Often, some guarantee of integrity is provided by lower level protocols. This is often in the form of error detection and correction. For example, Ethernet uses a cyclic redundancy check (CRC) to detect errors. TCP/IP uses checksums which will also aid in the detection of errors. Mobile communications often use forward error correction (FEC) codes that can both detect and correct errors in the received data. However, all of the above techniques have limitations in the number of errors they detect or correct, and none of them will protect against data that has been altered maliciously. This is because if the attacker can alter the data, they will most likely be able to change the CRC, checksum or FEC code.

The process of checking the integrity of a received packet is the same as that used for checking its authenticity, i.e., the use of a cryptographic hashing function and digital signatures.

8.2.3 Confidentiality

Confidentiality is the requirement on a system that only the intended recipient should be able to make use of the transmitted information. Confidentiality is a security requirement for a number of reasons. There could be a requirement to protect a financial transaction, the identity of the parties involved in the communication or for any number of other reasons.

The threat against confidentiality is eavesdropping. This is where a person or device can intercept or view traffic and then make use of it for surreptitious reasons.

The most common technique used to ensure confidentiality is encryption. There are a variety of different cryptographic algorithms offering different levels of protection including many open source algorithms as well as propriety implementations. For a long time, modern encryption techniques were impractical because there was no way to exchange keys securely. This is no longer the case as key exchange mechanisms such as public-private key cryptography, and techniques such as the Diffie-Hellman algorithm have been developed.

8.2.4 Availability

The availability of a system is the probability that the system will be operational at the instance it is required. Most modern communication systems demand a very high availability. Availability is often quoted as a number of nines, e.g., 3 nines is 99.9% (which corresponds to less than 9 hours 'down time' per year).

There are a number of reasons why a system may be unavailable which have absolutely nothing to do with security. For instance, a device might be out of range or there might be too many users attempting to access the system at a single moment in time. The system may also be made intentionally unavailable in order to carry out routine maintenance such as upgrading the hardware and/or software.

There are, however, reasons why availability might become a security issue. The most important example is when an attacker attempts to disrupt the normal operation of a communications channel. This is often called a denial of service (DOS) attack.

UNCLASSIFIED

DOS attacks are a significant problem on the modern-day Internet, and account for a significant percentage of all Internet traffic. DOS attacks can come in a variety of forms, ranging from a manual attack by an individual to massive, coordinated transmissions that often result from computer viruses. Although the motivation for these types of an attack is often a mystery to security analysts, they account for the vast majority of security related incidents. The main problem on the Internet is that these attacks often require very little skill as it is possible to use automated programs that are readily downloadable from a variety of hacking sites. These “off-the-shelf” attacks often prove to be extremely effective.

The vast majority of the attacks that take place on the Internet exploit some kind of bug in the targeted software. In theory, these attacks could be prevented by good software design procedures. However, in practice, the complexity of the software involved is so immense that software bugs will be prevalent for the foreseeable future.

Considering wireless systems, there are security issues affecting availability which potentially have no solution. One of the most obvious issues is that of intentionally jamming a system by saturating a given portion of the radio spectrum with a signal that blocks legitimate ones.

Potentially, SDR has both advantages and disadvantages when considering this type of attack. It may be more vulnerable because it may be possible to disrupt a small amount of key control traffic with wide ranging implications, or it may be possible to prevent software updates from being downloaded. However, SDR might offer better resilience to this kind of attack as it could be reprogrammed to use a different frequency or use a more resilient wave form.

There is no reason to believe that, given the opportunity, the same kind of attacks that plague the Internet would not be conducted against SDR. Therefore, strict measures need to be taken to avoid SDR from suffering from the same vulnerabilities that plague the Internet. (In Section 8.4 we will discuss the reasons why the general purpose architecture of a PC is fundamentally insecure and may not, therefore, be appropriate for an SDR architecture).

8.2.5 Summary

The issues raised can be simply illustrated by imaging the situation when an attacker, say Alex intends to undermine communications between Ulrike (the authorised transmitter) and Fabian (the authorised receiver); whilst Alex can mount various attacks of increasing levels of sophistication, e.g. from a simple denial of service, such as jamming a transmission, to something more complicated, such as a ‘replay attack’, Ulrike has at her disposal a choice of varying forms of encryption, authentication and tamper-proofing to provide a given level of protection. In the context of protection, issues requiring consideration include cost, level of system complexity, mechanisms for key management, etc. The four general requirements for security are summarised in Table 8-1, which also identifies the technologies that could be used to satisfy these requirements.

UNCLASSIFIED

Requirement	Technology	Implications for Design and/or Infrastructure
Confidentiality	Cryptography.	Extra hardware/processing power required to implement cryptographic algorithms.
Authentication	<p><i>Preshared Keys</i></p> <p>Techniques such as challenge handshake authentication protocol (CHAP) exist where a device can be authenticated securely if a preshared key exists.</p> <p><i>Digital Signatures</i></p> <p>Using public private key cryptography it is possible to digitally sign information which binds that information to a particular private key.</p>	<p>Administration of preshared keys in even a medium sized network may be impractical.</p> <p>A trusted third party is used to create a digital signature that binds a public key to an identity. This leads to the creation of certification authorities (CAs) and a public key infrastructure (PKI).</p>
Integrity	The same techniques that are used to ensure the authentication of the data also provides strong integrity checking.	The implications for the provision of integrity checking are the same as that for authentication.
Availability	<p><i>Robust Design Practices</i></p> <p>A more fundamentally secure architecture.</p>	Increased cost in software and hardware design.

Table 8-1: Summary of the technologies available to satisfy the four general security requirements of SDR

8.3 Security Analysis of an SDR Environment

In this section we look at security issues that are unique to SDR and suggest measures that might be needed to protect communication between SDR-enabled devices. We begin by describing the threat environment in terms of the possible types of attackers and the various goals that they might have. We then consider the types of traffic that might be present in an SDR environment and examine the security requirements for each.

UNCLASSIFIED

8.3.1 The Threat Environment

There has been much work done on analysing the security environment in which SDR will have to operate. Most of these analyses produce what is called a ‘threat vector’, that looks at security issues from several perspectives. One of the most useful examples of a threat vector for SDR was submitted as a paper to the SDR Forum [266]. This paper looks at the types of perpetrators and the different types of security violations that they could potentially cause. The following two sections summarise the findings of this paper by including the categories of perpetrators and types of security violations that were identified.

8.3.1.1 Perpetrators

A perpetrator is an individual who takes some action that may result in a security violation.

Negligent

Negligent covers those individuals that are not deliberately causing security problems but for one reason or another does so. Two types of potentially negligent users are identified. The first is the ‘normal’ user that may inadvertently overload the system, either by the sheer volume of other users or by requesting too much bandwidth. The second type of negligent user is the accidental interferer. An accidental interferer is a person who misuses equipment in a way that inadvertently interferes with other communication system users.

Unauthorised

The existence of a system attracts certain kinds of people who attempt to access the system beyond their normal access rights. Reference [266] defines three types of unauthorised users. These are ‘interceptors’, ‘probers’ and ‘impersonators’. An interceptor is a person that wants access to information to which they have no right. A prober is an individual who tests the system for weak points, perhaps just to see how it works. An impersonator is an individual who attempts to access the network using normal procedures but with false credentials.

Malicious

These individuals intend to steal service or information, or to disrupt system operation. Again, Reference [266] defines three types of malicious perpetrators, namely, ‘thieves’, ‘intentional interferers’ and ‘insiders’. A thief is defined as someone who wants to avail themselves of services or content that is offered for a fee without paying. An intentional interferer is a person with intent to disrupt or deny other users’ communications. An insider is an individual authorised to access some part of the system who misuses his/her access to commit unauthorised acts.

8.3.1.2 Security Violations

Security Violations are split into categories of action taken by the perpetrator.

Impersonation Violations

Reference [266] defines three ways that impersonation could be used to violate security. The first is when a user attempts to impersonate another user by either stealing or guessing his/her credentials. The second is where an individual attempts to impersonate the base station side of the communication link with the hope that legitimate users will be tricked into logging on to it. The third type is called a “man in the middle attack”, where a user attempts to intercept and then relay a legitimate user’s traffic to a genuine base station. This would enable them to view and/or modify the traffic.

Unauthorised Access

Unauthorised access to a system results when a perpetrator finds some way of getting around the normal security procedures of a system by a method other than impersonation. Several types of unauthorised access are identified by Reference [266]. These include extraction of the content of a transmission, unauthorised access to the control feature of a system and unauthorised access to keys and cryptographic material.

Denial of Access

Reference [266] identifies two possible ways in which a perpetrator could deny access to the system. Either they could disrupt the system as a whole to deny access to the user community, or they could find a way of altering parameters in such a way that one or more users gain elevated privileges at the expense of normal users.

8.3.2 Types of Traffic in an SDR Environment

Before the security requirements can be analysed, it is important to define what types of traffic will pass over the network. This will help to avoid a “one size fits all” approach that might result in a suboptimum solution [267]. If all traffic were treated equally, the worst case scenario would be chosen and all traffic would be given the highest possible level of protection. This would be a very resource intensive approach.

A more sensible approach would be to consider the different types of traffic and provide a range of solutions depending on the individual security requirements. The traffic can be divided into three categories.

8.3.2.1 Software Update

A software update is the term used for the transfer of data from one side of the communications link to the other. The update may consist of more than one packet and it is a reasonable assumption that the update cannot be used until the entire update has been received. This is effectively a file transfer.

UNCLASSIFIED

Examples of such software might include:

- Update of parameter tables for handoff algorithms
- Update of the state machine
- New radio air interface
- New operating system and drivers
- New modulation/demodulation scheme.

8.3.2.2 Control Traffic

Control traffic can be thought of as any information that is transferred in order to keep the current communications link open and operational. Control traffic will typically consist of a single packet of data. The most fundamental requirement for such traffic is latency. Control traffic is likely to be valid for a very short period as the environment in which the system is operating can change very rapidly.

Examples of control traffic might include:

- New transmit power
- Instruction to change frequency
- Instruction to change modulation scheme
- Instruction to change FEC code
- Request for the retransmission of data.

8.3.2.3 User Data

User data are the data that the user wishes to transfer over the link. As SDR is adopted, it is likely that there will be an ever increasing number of user data types. Unfortunately, the different types of traffic typically have different requirements in terms of bandwidth, bit error rate and latency.

Examples of user data include:

- Voice
- File transfer
- Multimedia (e.g., video) streaming
- Web traffic (i.e., the Internet).

8.3.3 The Requirements of a Secure SDR System

When considering the solutions for a set of security related problems, it is very important that a cost-benefit analysis is conducted [268]. This has to be done at two levels. One is from the perspective of the attacker who is attempting to breach the security. The other is from the perspective of the service provider who has to create and implement the required security measures.

UNCLASSIFIED

From the attacker's side it could be argued that any rational attacker will only attack a system if the information that they could acquire from such an attack is worth more than the effort that would be required to obtain it. It should be noted that the effort required generally decreases with time as new techniques and greater processing power becomes available.

From the service provider's point of view, the same rules apply. There is little motivation to provide security to the system at a cost that is greater than the cost that would result in a breach in that system's security. It is difficult to accurately evaluate the cost of such a compromise as there may be few direct costs. However, such costs that do exist are likely to be in the form of lost customer confidence in the system which might result in reduced use and hence lower sales revenue. Well informed users are also capable of performing a similar analysis of the protection provided by the system and of the information they wish to transfer over it.

8.3.3.1 Security Requirements for the Software Updates

An area that is very important for SDR is software security related to regulatory issues. A conventional radio is designed and then tested against a specific set of regulations. These regulations specify the maximum amount of power that can be transmitted in a given channel as well as the amount of residual power they can transmit into adjacent bands, i.e., out-of-band emissions. Without this external regulation, there could be significant problems with interference and incompatibility in every communications system.

The problem that is introduced with SDR is that these vital characteristics of a system can change as each new piece of software is downloaded and installed. In order to ensure that the new hardware/software combination conforms to the appropriate regulations, it is very important that only approved, compatible software is installed on a given piece of hardware.

In terms of security, there is a minimum of two requirements. These are authentication and integrity. The end user or device has to be confident that the new software will operate correctly and is appropriate to install. Conversely, the software provider has to be sure that the end user is eligible for the download and is using a compatible piece of hardware.

This suggests a two stage process. First, users have to authenticate themselves and the hardware that they are using to the software server. Thus, two digital signatures will have to be transferred. One is a digital signature that was installed in the hardware when it was manufactured to show that it is genuine. The other is the digital signature of the user which has been signed by the service provider. This will enable the software server to confirm that the user is entitled to the download and has the appropriate hardware to run it on.

The second stage is executed after the new software has been downloaded. This is the authenticity/integrity check of the new download by the user. This process can also be done by the validation of the digital signature received with the data. A crucial question at this stage is who will have to sign the new software. The number of signatures and the parties who will do the signing is a subject of much debate. Reference [269] suggests up to three signatures depending on the class of software.

UNCLASSIFIED

One signature could be from the originator of the software, which could be a third party vendor or the manufacturer of the terminal. The second is the network operator, who signs the download as an indicator of his/her approval for the use of the package on his/her system. The third signature could be from the regulatory authority in whose area the terminal receiving the download is operating.

If it is possible to ascertain that the new piece of software has been approved by the regulatory authorities and has not been altered since then, it can be assumed with some confidence that the new hardware/software combination is compliant with the current regulations.

In addition to integrity and authenticity, the service provider is likely to require that the software that was transmitted remains confidential to the user. To support this requirement the use of encryption to protect the data will most likely be used. This makes it much more difficult for someone to obtain the software by intercepting the transmission of the file. Using some form of link encryption will also make replay attacks (in the form of a fake server transmitting valid software updates) more difficult. Note that stopping the user from copying it and giving it to a third party is likely to be a greater challenge.

Table 8-2 summarises the requirements and possible solutions used for securing new software downloads.

Requirement	Justification	Solution
Authenticate the user	The software download may be a chargeable service.	User authenticates the request by digitally signing the request.
Authenticate the hardware	The software may only be compatible with certain hardware platforms for technical and/or regulatory reasons.	Transmit digital signature installed in the hardware during production.
Authenticate the download	The user needs to be confident that the software is genuine and security vulnerabilities have not been introduced by a third party.	Digitally sign the software with one or more signatures (e.g., software vendor, regulatory authority and/or service provider).
Confidentiality	To prevent any potential eavesdroppers obtaining the software free of charge.	Encrypt the software transmission process.

Table 8-2: Summary of the requirements for software download

8.3.3.2 Security Requirements for the Control Traffic

It can be seen from Section 8.3.2 that SDR will include a certain amount of control traffic. This is not dissimilar to current systems that also generate significant amounts of control traffic. For example, GSM and third generation CDMA systems implement transmit power control mechanisms whereby the mobile and the base station can instruct the other to either reduce or increase their transmit power. Control traffic is also used to facilitate cell handovers as mobile users move away from one base station and towards another. It is likely that SDR will have additional control traffic which handles aspects unique to SDR. For example, it may include instructions to change to a different waveform or frequency in order to improve the quality of service or to reduce congestion in a particular frequency band. Concepts such as dynamic spectrum allocation will also introduce requirements for additional control traffic.

As with GSM, it will be important for SDR to ‘protect’ control traffic. This requirement exists for a number of reasons. For example, the service provider will want to protect the billing information that could be derived from the control traffic. Furthermore, end users will expect that his/her identity be kept confidential. In more general terms, it is likely that if an attacker could easily read the control traffic, they would be able to mount a successful attack on the system with greater ease.

An example showing the benefits of providing security for control traffic concerns the robustness of the system to interference. SDR has the potential to make a communication system more resilient to both intentional and unintentional jamming if it has the ability to change to a different operating frequency on demand.

Unintentional jamming might result from a number of different operating scenarios. For example, it could simply be that there are currently too many devices trying to communicate in a given frequency band. Alternatively, a piece of faulty equipment transmitting at the wrong frequency might cause unexpected interference.

Although intentional jamming is rare in commercial situations, there are scenarios in which it has occurred. For example, in Japan, June 2001, many Internet-enabled mobile phones were affected by an email which caused the phone to dial Japan’s equivalent of ‘999’ [270]. This caused widespread disruption. In military situations, it is common practice to attempt to interfere with the enemies’ communications using a jamming signal. In this situation, switching to a different frequency or waveform might help if the enemy is only jamming selected frequency bands. Of course, making such a change would be futile if the enemy could detect what the new frequency/waveform was and modify their attack accordingly. Frequency hopping radios attempt to stay one step ahead by transmitting on a particular frequency for only a short ‘dwell’ time before changing to a new, predetermined frequency. As an aside, frequency hopping radios also help protect the transmitted content from eavesdroppers; without prior knowledge of the hopping ‘pattern’, significantly greater effort is required on the part of hostile third parties in order to ‘capture’ the transmitted message in its entirety.

The possibility that an attacker might attempt to surreptitiously inject spoof control traffic also has to be considered. If an attacker can instruct a system to change its power level or move to a different frequency band the whole system will quickly become unusable.

UNCLASSIFIED

The arguments above demonstrate that if system availability is to be maintained then authentication, integrity and confidentiality are all crucial requirements for the control traffic. However, the protection provided by the system might only need to be short term because any information discovered by an attacker will only be useful for a relatively short time period after transmission.

8.3.3.3 Security Requirements for the User Data

In a wireless environment such as an SDR scenario, there are two potential methods of protecting the user data. The first is to use end-to-end security, where the data are encrypted at the source and are not decrypted until they arrive at their final destination. The second involves securing the air interface itself, which is perhaps the weakest link in the data path.

At this point it is useful to consider the security that is provided by existing technologies that might ultimately end up competing with new SDR systems. Key examples are of such technologies are GSM and PSTN. GSM is the dominant second generation mobile phone technology in Europe and PSTN refers to conventional land line telephones.

In GSM, only the over-the-air traffic data are encrypted. Once the frames have been received by the base station, it decrypts them and sends them in plaintext over the operator's backbone network.

The original intention when the GSM standard was being developed was to keep the security mechanisms that would protect the air interface a secret. The idea was that if the protocols were released to hardware/software vendors on a strict need-to-know basis security would be enhanced. This is effectively security through obscurity and should not be regarded as a sensible strategy. Inevitably, the standards slowly leaked into public domain and several potential weaknesses were highlighted. A more sensible approach is to publish open standards and rely on the strength of the cryptographic algorithms and their implementations for security.

PSTNs often provide no security in the form of data encryption at all. However, we note that circuit-switched networks are generally regarded as fundamentally more secure than packet-switched ones. This is because the data path in a circuit-switched network is much more predictable than that in packet-switched network (such as IP on the Internet). This makes it more difficult for an attacker to intercept the data and "tap the wire". Therefore some measure of security is provided at the physical layer reducing the need to add security features at a higher layer. We further note that a physical 'tap' is required in order to eavesdrop on PSTN traffic; in a wireless network it is much easier to 'listen in' without being detected or leaving any evidence.

The PSTN was originally developed for voice traffic but it is now commonly used for data traffic. It is also quite common for end users to routinely pass extremely sensitive data over such links. Common examples are online shopping where credit card details are revealed, Internet banking where extremely confidential data are transferred and home working using a VPN dial up connection where commercially sensitive information are accessed.

The reason that the above transactions can be done securely is because security is provided at a higher layer in the protocol stack. In the examples above, security is added at the IP layer (i.e., IPSec) or at an even higher level such as with a secure sockets layer (SSL).

UNCLASSIFIED

It is sensible to use this layered approach when deciding on what level of security has to be provided by SDR. One approach might be to provide a minimal level of security at the low-level SDR layers and rely on the higher level layers to provide security for traffic that requires enhanced levels of security. Although this might lead to encrypting traffic more than once, it will considerably simplify the design and improve compatibility with other systems/technologies.

A crucial consideration that also gives merit to the layered security model is that the end user is unlikely to want to place a significant level of trust in the provider's network. This means that end-to-end encryption is likely to be a requirement at higher levels anyway.

8.3.4 Summary

All three types of traffic have a requirement for a certain level of confidentiality, integrity and authentication. The benefits of a layered security model have also been discussed.

An appropriate way forward seems to be to provide a minimal level of authentication, integrity and confidentiality by default at the lowest levels of the protocol stack. When there is a requirement for a greater level of protection, e.g., for services transferring particularly sensitive data, then this should be provided at a higher level. An example of how this approach is used on the Internet was described.

The fundamental security layer requires standardisation before SDR can be deployed on a large scale. However, a key requirement when considering standardisation of the lower level protocols for SDR is that the ability to upgrade to incorporate new security features as they become available should be supported. The exact details of who should digitally sign software downloads and the format of any digital signatures also has to be standardised in order to meet regulatory requirements and ensure compatibility.

Table 8-3 summarises the requirements for each type of traffic.

Type of Traffic	Requirement for Confidentiality	Requirement for Integrity	Requirement for Authentication
Software download	Medium	Strong	Strong
Control traffic	Medium/strong (Short term only)	Medium	Medium
User traffic	Medium (User data may have additional security enhancements added at higher levels of the protocol stack)	Low	Low

Table 8-3: Summary of the security requirements for different types of SDR traffic

8.4 Hardware Security

In the previous section we discussed security issues in the context of SDR. We now move to consider why the architecture of a general purpose processor may not be appropriate for SDR. First we explain the deficiencies of a general purpose processor by including an example of a particular type of vulnerability, buffer overflow. We then give an outline of possible solutions to this problem.

8.4.1 The Insecurity of a General Purpose Processor

The architecture of a modern PC is not very well suited to providing security for its software. One of the primary reasons for this is the structure of its memory. At any given time, there are potentially hundreds of programs running which all have access to a single, shared pool of memory.

Under normal conditions this simple architecture works extremely well and everything runs efficiently. However, under certain circumstances, programs can be made to inadvertently read or write to memory ‘belonging’ to another process. This has a range of implications. One of the most crucial parts of the memory is known as the stack. The stack is a contiguous block of memory.

All modern programming languages use the concept of functions which are used to structure programs. A function can be thought of as a mini program which is launched from inside a larger one. When a function is called, certain data are ‘pushed’ onto the stack, these data include any variables that are used while the function is running and a ‘return’ address. The return address is the location of the next instruction after the original function call.

Below is the pseudo code for a trivial program implemented in C. It is not crucial for the reader to have knowledge of the C programming language to understand the following explanation.

```
void function()
{
    char buffer1[255];
    char buffer2[255];
    :
}

main()
{
    :
    function();
    :
}
```

When *function()* is called from within the main function, the stack would look similar to that shown below.

UNCLASSIFIED

[buffer2 (255 bytes)][buffer 1 (255 bytes)][return address]

After *function()* has finished executing, the data will be removed from the stack and the original thread will continue to be executed. The problem is that the programmer may not have provided any mechanism to ensure that the size of the buffer does not exceed the size allocated to it. For example, suppose *buffer2* is memory that has been made available to accept an input from a user in the form of a string, perhaps his/her name.

Under normal circumstances 255 characters would be more than adequate to hold any conceivable name. However, it is important to know what would happen if more than 255 characters were input. In this case, as soon as the 255 bytes allocated to *buffer2* had been used up, the program would start to write to the space allocated to *buffer1*, corrupting the data stored in it. After *buffer1* had been filled, it would then write over the return address.

Once *function()* has finished executing, the processor will attempt to load the instruction that is being pointed to by the return address. However, the return address has just been corrupted. Under normal circumstances, this would most likely lead to a crash of some sort. Unfortunately, it is possible for an attacker to manipulate the input to execute arbitrary code.

Suppose that instead of entering his/her first name, he/she entered the ASCII representation of a small executable program. They could then potentially cause the return address to be over written in such a way that the return address points to the start of *buffer2*. The program that the user has just entered as a string will now be executed. This program could do anything ranging from crashing the system to installing a Trojan program on the victim's machine.

This type of overflow mechanism accounts for a large proportion of the vulnerabilities in modern software. In this case, the problem could have been completely avoided if the program had checked that the data input by the user were less than 255 bytes in length before copying the string to memory. Indeed, this might be regarded as best practice coding anyway.

Unfortunately however, it is widely recognised that software accounts for some of the most complex systems mankind has ever developed. The ever increasing demand for additional functionality and features leads to systems where the absolute proof of correctness is not (commercially at least) practical. This leads to the unfortunate conclusion that unless the memory of system is managed or protected in some way, there will always be the possibility that one program can gain access to another program's data.

The problem of shared memory makes it difficult to build a secure system, where it is necessary to have different levels of access for different processes/tasks.

8.4.2 Technologies for Providing Memory Security

The problem with memory protection has been around for a long time and a number of solutions have been developed which could be incorporated into the design of an SDR architecture. This section looks at a number of these technologies and identifies advantages and disadvantages of each.

UNCLASSIFIED

8.4.2.1 Virtual Machines

The term ‘virtual machine’ (VM) is used in a number of slightly different contexts. A definition that is appropriate here is “a simulated computer in that it runs on a host computer but behaves as if it were a separate computer” [271]. Each virtual machine can be given its own set of resources that are isolated from those used by other virtual machines. A good example of such a technology is the virtual machine that is employed by the Java programming language. Java uses its programming language mechanisms to enforce secure use of memory. The Java virtual machine (JVM) enforces the Java language’s type safety, preventing programs from accessing memory or calling methods without authorization. Existing JVM implementations enforce a simple ‘sandbox’ security model that prohibits untrusted code from using any sensitive system services.

The biggest problem with using a virtual machine approach is that there are significant overheads in terms of memory usage and processing power. Virtual machines are an excellent way of providing security for most application level programs, but it is doubtful that they would be suitable for the high-speed DSP algorithms that would be employed in SDR.

8.4.2.2 Separate Hardware

It is possible that the virtual machine idea could be taken a step further by using physically separate hardware for each process or class of process. This is the strategy used when developing high grade cryptographic devices, where keeping different grades of traffic separate is absolutely essential. Each level of process will have its own CPU and memory. This provides a great deal of protection in the form of a ‘fail safe’ architecture. However, the cost in terms of additional hardware and design complexity is likely to preclude this type of solution.

8.4.2.3 Memory Control Module

One approach suggested in Reference [269] for memory management of SDR modules was to incorporate a ‘radio security module’ (RSM), which controls when a new program can be installed.

The RSM should have output control signals that enable or disable write access to memory for both program and user data. Thus, the RSM would only enable write access when the software to be installed has successfully passed all the required security screening mechanisms. Only then could a new software application be installed and activated to run.

This is a good approach for making sure that only authenticated programs get installed to memory. However, if these programs have programming errors, then this approach is not fail safe.

8.5 Conclusions

In this chapter we have looked at the security requirements of an SDR system. The types of traffic were identified and the requirements for each were analysed in terms of authentication, integrity, confidentiality and availability. Current technologies that could be used to meet these requirements were identified. There does not appear to be a technological gap between current capabilities and those required for a secure implementation of SDR. However, standardisation is a significant obstacle. It is essential that security algorithms and procedures are standardised to ensure current and future compatibility.

The specific problem of downloading software updates over an air interface was discussed. While there seems to be no technological barrier, standardisation is again an issue. Several digital signatures will be required for each piece of downloaded software in order to meet likely regulatory requirements. Exactly who would need to 'sign' software downloads needs to be standardised before the large scale deployment of over-the-air SDR updates can be realised.

Finally, we have discussed the security implications of a general purpose processor and explained why conventional, general purpose architecture may not be appropriate for an SDR device. Several technologies that may enhance the security architecture of an SDR terminal have been considered.

9 Radio Management and CR



By Daniel Bradford, QinetiQ.

9.1 Introduction

In this chapter we look at cognitive radio (CR) and how it could be used to make more effective use of the available radio spectrum. For the purposes of this section, we define a cognitive radio as a radio that is aware of its surroundings and can adapt its functionality to improve its performance in some way.

There are two main thrusts to this section. We start by looking at ‘radio management’. Radio management is about making as much use as possible of the limited radio spectrum available whilst ensuring that all systems can coexist without causing unacceptable inter-technology interference. We will look first at how the radio spectrum is currently managed and comment on the pros and cons of such a system.

Some measurements of typical spectrum utilisation are given which highlight how inefficiently spectrum is typically used. These observations provide the motivation for the extensive research that has been carried out in this area. Some of this research is introduced and the main results and conclusions are outlined.

Some simple calculations are introduced which give some theoretical limits on the gains that might be achieved through managing the radio spectrum in a more efficient manner. We then introduce some more realistic schemes from which more practical efficiency gains are given.

The second main thrust of this section is cognitive radio itself. The term cognitive radio was originally coined by Joseph Mitola in 1999. Since then, however, it has become somewhat overused and now has several slightly different meanings. Therefore we start our discussion on cognitive radio by discussing the different ways in which the term is used, and the implications that they may have for future radio systems.

As well as increased bandwidth efficiency, cognitive radio potentially offers several other advantages over more conventional radio systems. These ideas are discussed in detail and the possible gains that might be achievable are identified.

There are already some examples of systems that are starting to manage the radio spectrum using cognitive radio techniques. While these systems are quite primitive, they demonstrate the strong push for more efficient systems. Some example systems that demonstrate limited cognitive behaviour are identified.

We conclude this section with some observations on the technical and commercial barriers which may potentially stop a fully cognitive radio from being deployed. We note here that a key enabling technology for cognitive radio is software defined radio.

9.2**The Current Radio Management Scheme**

The current system of spectrum allocation is very static in nature. Each system is allocated a range of frequencies that can be used. All characteristics of the transmitted signal are specified including the maximum transmit power and the power that can be transmitted out-of-band. Usually, any system has to be accredited and licensed before it can be deployed. Once an allocation has been made, it is likely to remain in place for many years. Making any changes to allocations is a difficult and often prolonged process.

The current system offers several important advantages, which explains why it is currently used. The first and most important benefit is the simplicity that the system offers; if each system only transmits in its own band and within the specified power limits, all the services can coexist and the individual systems will work as intended. The static system also ensures that appropriate amounts of spectrum resource are allocated for public, government and public safety (i.e., emergency) use which means that no system of real-time traffic prioritisation is needed. The scheme also ensures some level of ‘fairness’, as it ensures that low power devices such as mobile phones are not swamped by more powerful devices that might otherwise attempt to use the same band.

There are, however, several short comings of the current spectrum allocation policies which have motivated the search for a better approach to spectrum management. The neXt Generation (XG) working group [272] has highlighted the two main deficiencies in the current system. The first is the inefficient use of the frequency spectrum that results from the current method. Reference [272] describes how, if the current trend is not changed, the entire ‘useful’ spectrum will be used. The main reason that the current system has been a viable option for so long is that advances in electronics have meant that increasingly higher frequencies can be used. However, very high frequencies have the propagation property of being increasingly line of sight and subject to greater path loss effects. These characteristics severely limit the range of applications that the new frequency allocations can be used for. Whilst very high frequencies might be suitable for high-bandwidth, fixed, point-to-point wireless links, they are less well suited for mobile applications.

The XG working group also highlights the difficulties of deploying such a system and comments on how extensive, frequency by frequency, system by system coordination is required for each country in which a system is operated. The working group also states that as the number, size, and complexity of operations increase, the time for deployment is becoming unacceptably long.

The working group published a figure, parts of which are reproduced in Figure 9-1, which shows that, whilst the spectrum allocation tables (Figure 9-1 (top)) indicate that the radio spectrum is very crowded, only a very small proportion of the usable spectrum is typically in use at any instance in time (Figure 9-1 (bottom)). In this case, it is quoted as being six percent, which is a fairly typical empirical result.

UNCLASSIFIED

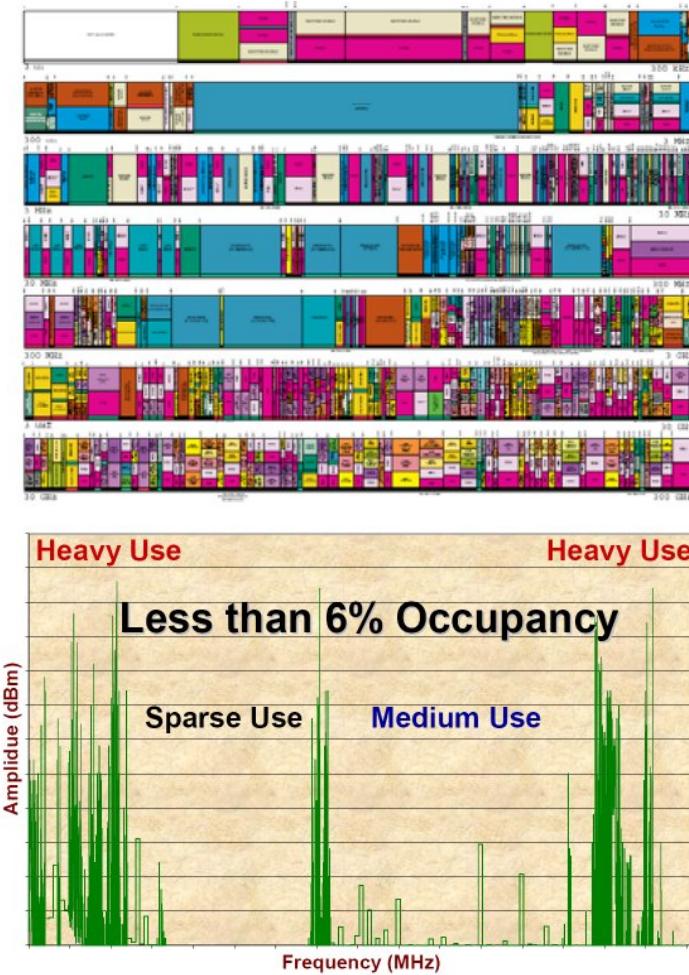


Figure 9-1: Snapshot of current spectrum allocation (top) and typical spectrum use (bottom) [272]

There are also examples of frequency allocations where a given frequency is only ever used in a few (some times in only one) geographic locations, for example, fixed satellite services. This clearly illustrates how a very valuable resource is often greatly underused.

It is important to note that the frequency allocation of a system is only one dimension of the whole radio management problem. Other dimensions include (but are not limited to) space, time and code. For example, a frequency that is being used in one location can be reused at another location provided that the other device is an acceptable distance away. This is called frequency reuse and is the reason that regulators specify a maximum transmit power (and therefore maximum range) for each frequency band or service. Note that, with the possible exception of CDMA network operators, cellular operators already rely on frequency planning techniques to optimally reuse their allocated RF carriers across the base stations in their networks.

9.3 Spectrum Trading

It is now widely accepted that in order to make better use of the precious radio spectrum, greater flexibility is required in the way it is used. A leading idea for increased flexibility is to introduce spectrum trading. This is where the primary (legacy) licence holder is allowed to sell the rights to use part of the spectrum to a third party.

Although this is widely regarded as a good idea, many companies and researchers have different ideas of how this is best done. This section includes a range of different ideas ranging from the long-term lease or even permanent transfers that have been proposed by Ofcom, to highly dynamic spectrum allocation, where a piece of spectrum is only leased for the instance that it is required. The remainder of this section outlines the various proposals that have been made.

It should be noted that current proposals would almost certainly require a change in regulations. The reason for this is that, under the current rules, it is not permitted to use the spectrum for any purpose other than for which it was originally licensed. This de-regularisation of spectrum usage is often referred to as spectrum liberalisation.

9.3.1 Permanent to Semi-Permanent Transfer of Spectrum Allocation

Ofcom has published consultation documents that describe how they may change their policies to allow primary (legacy) license holders to sell or lease part of their allocation to a third party. Ofcom cites its duty under the Communications Act of 2003 to ensure the optimal use of the radio spectrum to further the interests of citizens and consumers as the motivation for its new approach.

The key points of the consultation document are that Ofcom will have to sanction each trade. Ofcom will publish acceptable classes of services for each part of the radio spectrum and will act to prevent anti-competitive trading, which is not in the interest of consumers. The document also concludes that the current procedure for policing the spectrum, which involves investigating complaints and punishing violators, would remain adequate.

The feedback that was received from the consultation document was extensive and largely supportive. Whilst this proposal is the least flexible, as each transaction will have to be appropriately sanctioned, it could be very effective at increasing the utilisation of the radio spectrum.

9.3.2 Long-Term Dynamic Spectrum Allocation

As discussed in Section 9.2, there are often large portions of spectrum that are not used. In certain geographic locations, a piece of spectrum will often remain unused for long periods of time. Many researchers believe it may be possible for CR devices to reuse this spectrum on a non-interfering basis.

An idea that is often discussed in the US is to allow CR devices to make use of unused TV channels. Reference [273] gives an excellent explanation of why this idea would be relatively simple to implement and would open up huge 'chunks' of spectrum at a time.

UNCLASSIFIED

Each TV channel has a 6 MHz channel with a well defined centre frequency. The well defined channel spacing and the fact that TV signals get switched on and off only infrequently makes unused TV spectrum an excellent starting place for cognitive radio. Reference [273] summarises an analysis that concludes that a CR device could detect a TV signal at a sufficient range to allow low power transmissions that would not cause significant interference.

9.3.3 Short-Term Dynamic Spectrum Allocation

To increase the utilisation of spectrum yet further, many researches have proposed the idea of dynamic spectrum allocation (DSA). This takes the basic spectrum trading processes outlined in the previous sections a step further. The key point is that it is dynamic, which means that spectrum is only obtained as it is required and on a short term basis.

All the proposed schemes assume that there is a central ‘pool’ of spectrum that could be shared between two or more devices, services or networks. The proposals differ on how spectrum assignment should be carried out and exactly how dynamic it should be.

In all but the fully dynamic case (as discussed later in Section 9.3.3), SDR would not be a prerequisite to the proposed schemes. The reason for this is that the extra intelligence would only be needed in the base stations rather than the individual devices. Note that the implication here is that we are primarily considering client/server, e.g., cellular, networks.

9.3.3.1 The Use of Queuing Theory

Simple queuing theory can be used to show that dynamic spectrum assignment could yield significant gains in spectrum efficiency. Consider the case of a bank with people queuing to see a cashier. In this analogy each block of frequencies is represented by a cashier and each packet, connection or request for transmission is represented by a person in the queue.

The analogy of the static case would be represented by a separate cashier for each service. For example, one cashier for payments, one for withdrawals, one for foreign currency exchange and another for general enquiries. As the customers come in, they join the appropriate queue and wait to be served. This example is shown in Figure 9-2 (top). Clearly, this is not very efficient and will result in some cashiers sitting idle and some cashiers with long queues. This is exactly what was observed in Figure 9-1.

UNCLASSIFIED

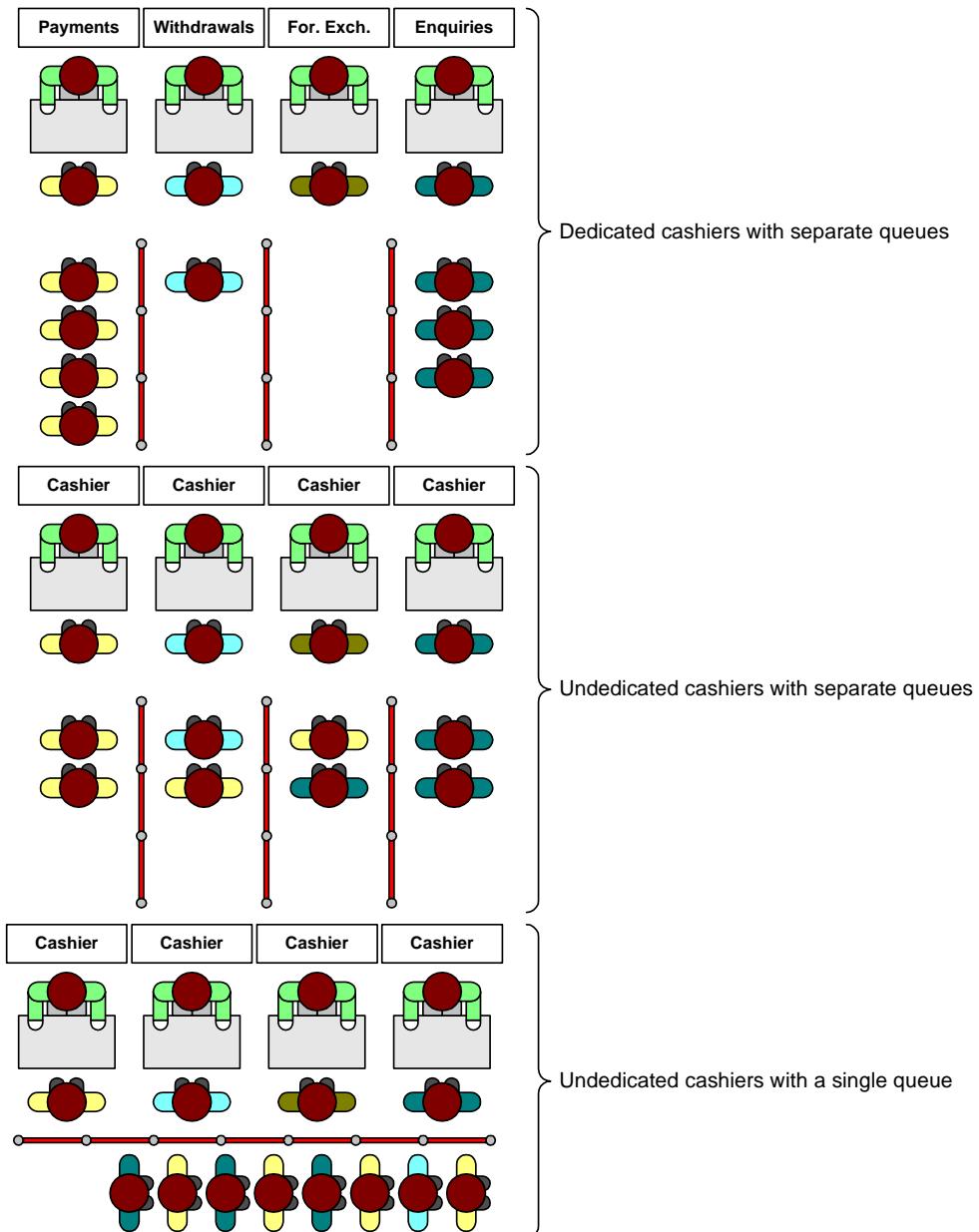


Figure 9-2: 'Bank cashier' analogy to the different approaches to spectrum trading

The next level of efficiency would be to train all of the cashiers in all of the required services but to still have multiple queues, as shown in Figure 9-2 (middle). This approach would mean that the queue lengths would be reduced and cashiers would be more highly utilised. Inevitably, however, some transactions will take longer than others. Thus, whilst the cashiers are used efficiently, the queues will move at different rates causing frustration to the customers.

UNCLASSIFIED

The highest level of efficiency and the one aimed for by researchers is to have a single queue for all of the checkouts, as shown in Figure 9-2 (bottom). As each request is dealt with by a cashier, the next customer is informed and they proceed to the appropriate counter. This is clearly more efficient than any other scheme (and is, incidentally, the scheme adopted by most modern banks). This analogy shows why several services sharing a single pool of spectrum are likely to offer significant gains in utilisation and efficiency.

9.3.3.2 Dynamic Spectrum Allocation at the Network Level

Dynamic Radio for IP-Service in Vehicular Environments (DRIVE) is a European organisation with the specific objective of improving network services for vehicles. In particular, they are interested in high-quality wireless IP services. Even though their remit is quite narrow, they have documented some interesting ideas for making more efficient use of the radio spectrum.

Their first idea outlines an approach in which the spectrum allocated to the radio network could be altered depending on the current demand. Leaves et al [274] talks about contiguous and fragmented spectrum allocations, and uses the concepts of radio access networks (RANs). In the most simple, contiguous case, each RAN is allocated a block of spectrum and each block is separated by an appropriate guard band. The size of this block is allowed to vary depending on the demand on the system.

This concept is shown in Figure 9-3. The fixed spectrum allocation (FSA) is the scheme that is currently employed, where each service is allocated a fixed block of spectrum. The scheme in the middle shows a scheme where the allocation to a RAN is allowed to shrink and grow depending on the current demand, and represents a contiguous scheme. The third scheme is the fragmented dynamic spectrum allocation approach and allows RANs to be allocated frequency blocks that are not contiguous.

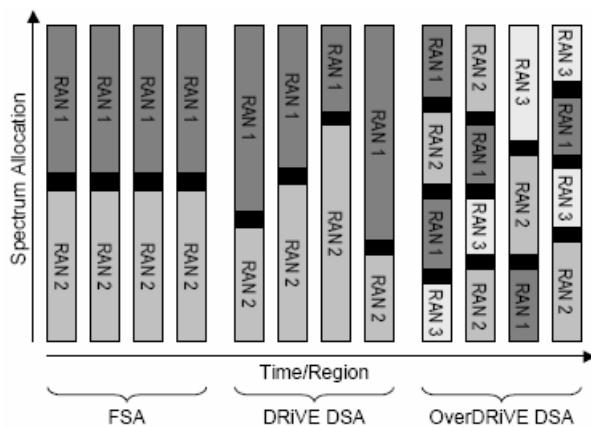


Figure 9-3: Spectrum sharing at a network level [274]

UNCLASSIFIED

This approach would undoubtedly give some gains in efficiency. The additional complexity for the mobile devices is likely to be relatively small. The devices would have to be designed to operate over larger frequency ranges which would become increasingly difficult as the frequency range was increased. The additional complexity of managing the frequency allocations would probably take place in the base stations.

DRIVE also gives an example of how the method could be used to increase the overall capacity of a system by exploiting the time diversity of different systems and services. The main example in [274] describes how two services, namely, UMTS (i.e., third generation mobile phones) and digital video broadcast-terrestrial (DVB-T), could be used to provide voice and multicast video streaming respectively.

Measurements have shown that the demand for these services changes throughout the day, as shown in Figure 9-4 [274]. The report also shows that the peak demand for these services is fairly uncorrelated, which means that the peak total demand for the two services is less than the sum of the peak demand for the two services. Figure 9-4 shows that, in this example, the peak of the summed curve is in fact 30% lower than the sum of the two peaks. Currently these two systems have separate frequency allocations. This means that no savings in bandwidth are currently realisable. However, if these two services were allowed to share the same pool of bandwidth, then the same peak capacity could be achieved with significantly less bandwidth.

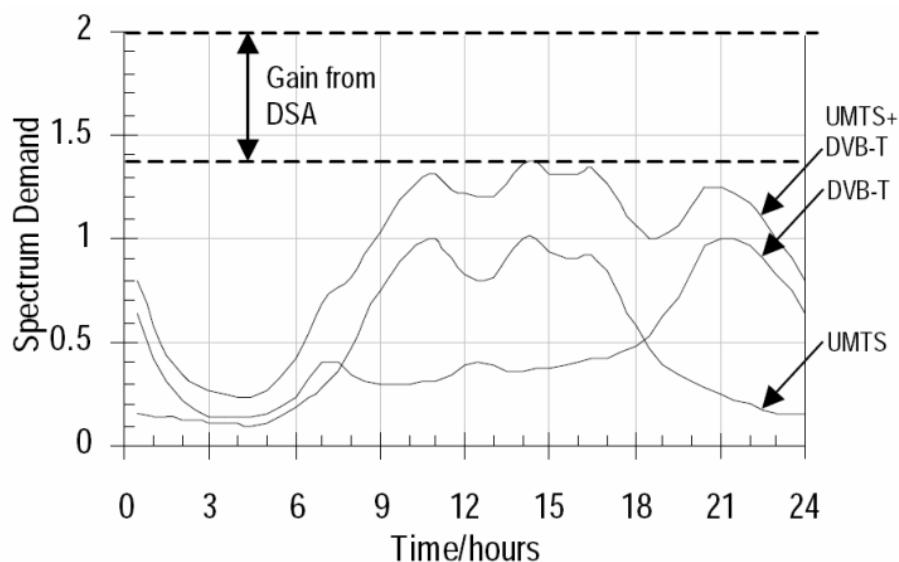


Figure 9-4: Example network usage for two services over a 24-hour period [274]

The report concludes that while a more realistic system may not produce such high gains in efficiency, there are still potential savings to be made. Note that, however, these savings would be at the expense of significant extra complexity in the system.

9.3.3.3 Fully Dynamic

The next logical step from the centrally controlled radio management scheme is to a decentralised or distributed scheme. In this scenario, each device would look for an available space in the spectrum and then transmit. This scenario is significantly more complex than the previously discussed scheme and would require a fairly complicated negotiation/multiple access procedure.

Figure 9-5 shows how a cognitive radio could be implemented using three distinct 'blocks'. These are the software radio block, the spectrum monitoring and options block and the policy box block.

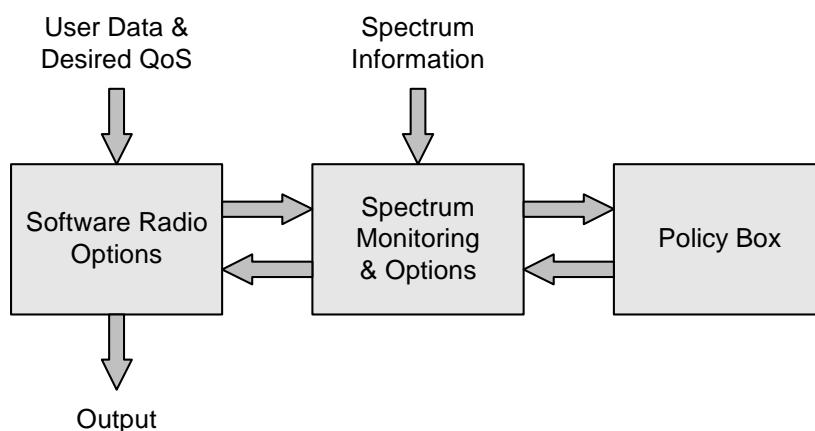


Figure 9-5: A block diagram of a fully cognitive system

The first block is the software radio element. The user data and the desired QoS provide the inputs to this block. The software radio then runs an algorithm to determine which waveforms it could use to meet the user's requirements. A list of the available options is then fed into the second block, the spectrum monitoring and options block.

The spectrum monitoring and options block has access to which parts of the radio spectrum are currently in use and which parts are not. It can use this information, along with the list of options from the first stage to rule out certain waveforms. For example, a particular service may not be available in the current location or it may be overloaded. This information is then fed into the policy box which evaluates the available options and the associated consequences of using any of them.

It is the policy box that will contain the intelligence of the system and accrue knowledge by remembering previous experiences. The policy box will also have to consider the current regulatory requirements to determine whether the chosen waveform is actually legal in the current location. Note that, however, unless the software radio provides a suitably large number of options and the spectrum monitoring block can give detailed and accurate information on the current environment then no sensible decisions could be derived in the policy box, no matter how sophisticated it is.

UNCLASSIFIED

This section has described a fairly sophisticated idea that has not yet been implemented. All three elements of the system require technology that is fairly new and not widely used in industry.

9.4 Cognitive Radio

Cognitive radio is seen as the key enabling technology that will allow dynamic spectrum trading. The reason for this is that, in many of the implementations, the radio will have to be aware of the current spectrum usage in order to make informed decisions about what spectrum it should attempt to use. In a dynamic environment, the devices will most likely have to also be aware of other devices, and have strategies to enable them to coexist.

This section describes the various definitions and uses of the term cognitive radio and describes how using cognitive approaches offers significant improvements over existing radio systems. While the idea of a fully cognitive radio is fairly new, technologies currently exist that could be viewed as cognitive to a greater or lesser degree. Some examples of these technologies are included in the final part of this section.

9.4.1 Definitions of Cognitive Radio

As stated in the introduction to this section, the phrase “cognitive radio” has become an overused term to describe significantly different ideas and technologies. Almost all definitions of the word “cognitive” describe the behaviour of humans. However, a definition that could apply equally well to a machine is “the mental process of knowing, including aspects such as awareness, perception, reasoning, and judgment”.

A definition of cognitive radio produced by the IEEE [275] is as follows:

“A cognitive radio is a radio frequency transmitter/receiver that is designed to intelligently detect whether a particular segment of the radio spectrum is currently in use, and to jump into (and out of, as necessary) the temporarily-unused spectrum very rapidly, without interfering with the transmissions of other authorized users.”

While the radio described in this example is undoubtedly cognitive, it precludes the use of the term cognitive radio for any purpose other than spectrum management.

The lowest common denominator for any definition of a cognitive system is self awareness coupled with an ability to change its behaviour depending on the knowledge accumulated via this self awareness. The reason that the term has become overused is that a system can, potentially, be aware at several different levels of the radio’s communication stack, and can be used to achieve several different goals.

An example of a cognitive radio at the highest level was given by Joseph Mitola [276]. The example describes how the incoming and outgoing multimedia content is parsed in order to gain information. The example goes onto describe how a radio could learn that the user of the radio has ordered a taxi and is, therefore, planning to become mobile and presumably change the communication requirements of the system. This is, perhaps, a very extreme form of cognitive radio and probably not realisable with current technology.

UNCLASSIFIED

A radio can also be cognitive at a much lower level. For example, if a radio can adapt its output power depending upon the current circumstances then it is cognitive. This is fairly straight forward but offers immediate advantages in terms of battery life and frequency reuse. Note that open and closed loop power control mechanisms are already an integral part of modern CDMA systems.

9.4.2 Possibilities for Cognitive Radio

Reference [275] expands the definition of cognitive radio to one that automatically finds and accesses unused spectrum across different networks. It also states that the goal of a cognitive radio is optimisation, i.e., to find the best link. However, there are multiple ways in which the best link could be calculated. The link metric is likely to be a function of cost per unit throughput, latency, BER, bit rate, power requirements plus several other attributes. Crucially, the metric calculation is likely to be different for each type of traffic the device is transmitting or receiving. For example, voice traffic would have different requirements to a file transfer. After the radio had assimilated all the necessary information it will have the ability to reconfigure itself to create the best possible link.

As well as finding the best link when the session is established, the radio would continuously adapt so that it could seamlessly roam across networks, always maintaining the best possible link. This means that the radio needs to be fully aware of its environment in terms of what spectrum is currently in use, what services are currently available and its geographic location.

While a conventional radio may be able to use a relatively simple multiple access method, such as frequency division, a cognitive radio can be a lot smarter in the way that it coexists with other systems. For example, the National Telecommunications and Information Administration (NTIA) at the US Department of Commerce has proposed an 'electrospace' model to describe the ways that today's wireless systems can coexist. They propose the following variables; physical location (latitude, longitude, and height); frequency; time; and direction of arrival (azimuth and elevation) [277]. Many researchers argue that even the electrospace framework is still too limited. For example, it does not include techniques such as low-power underlay, cooperative mesh networking or other modern ideas to increase spectrum capacity [278].

A very simple and effective way that a cognitive radio could adapt to its environment is to adapt its modulation scheme. Figure 9-6 shows a very simple example of how a radio could improve its throughput by switching to a more bandwidth efficient scheme when a high signal level is available and to a more noise tolerant one when the signal to noise ratio is poor [279].

UNCLASSIFIED

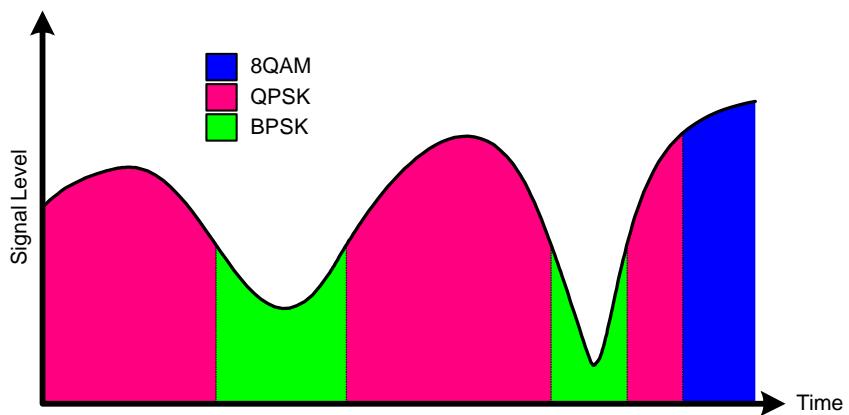


Figure 9-6: Adaptive modulation

As well as switching between these generic waveforms, it has been proposed that a cognitive radio could be used to access several different standard waveforms or protocols. An example of this is shown in Figure 9-7. This example is also taken from Reference [279], which describes a child on a journey playing a wireless computer game. As the device moves away from the Wi-Fi hot spot, the radio establishes a 3G connection and seamlessly continues the call. The device also switches to GPRS when the 3G service is not available. As better communication options become available, the radio will switch back again.

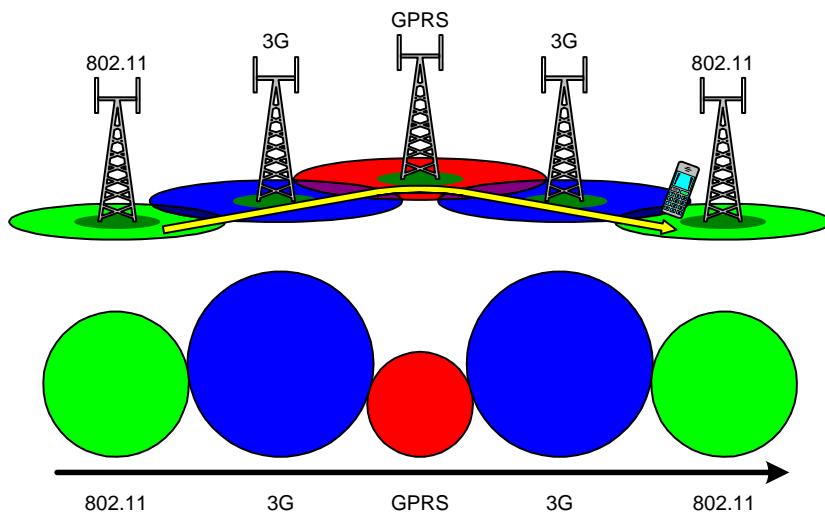


Figure 9-7: An example of a cognitive radio selecting the 'optimum' waveform throughout a journey

UNCLASSIFIED

Although the hypothetical device is cognitive in the respect that it is aware of the services that are available and is able to make decisions based on this information, it is not much more than a simple switch. The radio would only require a very simple algorithm as it would always favour the 802.11 over anything else and would always favour 3G over GPRS. That said, it would probably be a very popular and useful product if were it available.

9.4.3 Examples of Primitive Cognitive Life Forms

Although the research into advanced cognitive radios is a relatively recent occurrence, there are many current examples of systems that have some degree of cognition. These examples show that with just a small amount of intelligence in the device, huge improvements in system performance are possible. These improvements range from an increase in capacity to being able to exist at all.

9.4.3.1 Voice Activity Detection and Discontinuous Transmission

For a long time, techniques have been used to reduce the bandwidth used during international telephone calls. The most common technique is to not transmit data when there is no voice activity on the line. Extensive research has been done on conversation modelling but, very simplistically, a participant in a two-way conversation talks for about half of the time. This potentially allows for a 50% reduction in bandwidth utilisation.

More recently, similar techniques have been used in conventional mobile cellular telephone systems for exactly the same purpose. Here, the mobile station only transmits when there are data waiting to be sent. This feature, generally known as discontinuous transmission (DTX), not only reduces interference to other users, it also implies reduced power consumption and, therefore, increased battery life.

9.4.3.2 Passive Scanning

Often a mobile device will have several channels over which it can attempt to transmit its data. While earlier systems required the end user to manually switch over to an alternate frequency, more modern systems are now emerging that can passively scan for other users and select a channel accordingly.

The FCC identified some examples of such technologies in their cognitive radio Notice of Proposed Rulemaking (NPRM). They referenced cordless telephones in the 43.71 to 44.49 MHz band that avoid channels occupied by private land mobile systems by using passive scanning techniques to automatically select unoccupied channels [280]. Here the device is improving its performance by reducing possible interference. Another example system that uses dynamic channel allocation is the DECT system [281].

The FCC also cite U-NII band devices operating in the 5.25 to 5.35 GHz and 5.47 to 5.725 GHz bands that are required to use similar techniques to avoid interference with US Federal Government operations. This is an example of a device that may never have been licensed unless this capability had been utilised.

9.4.3.3 Protocol Switching

One of the most common uses of the term cognitive radio is to describe a device that can switch between different mobile services. Examples of services that a mobile device might potentially be interested in using include GPRS, Bluetooth wireless technology and Wi-Fi (i.e., 802.11).

An example of a system being developed that is capable of switching between two services is BT's Bluephone [282]. This is a mobile phone that can switch between BT's fixed broadband network and BT's existing mobile network. Typically the phone would use Bluetooth wireless technology to connect to a wireless access point when one is available in order to reduce the cost of expensive mobile calls.

9.4.3.4 Automatic Gain Control and Transmit Power Control

Many modern systems are capable of determining the distance from their intended receiver or the presence of a nearby system and adjusting their output power accordingly. This has the effect of reducing unnecessary interference with other systems and therefore increasing the possibility of frequency reuse.

A prominent system that employs this technique of automatic gain control is GSM. A GSM device (a base station or mobile station) is capable of measuring the power of the signal it receives and feeding back information to alter the transmit power. As well as reducing inter-cellular interference, it also enables the mobile unit to preserve power which greatly increases the battery life of the device. All of these characteristics are highly desirable in a wireless mobile environment. Note that open and closed power control loops are also an integral part of modern CDMA systems, e.g., IS-95 and 3G W-CDMA systems.

A good example of another system that also adjusts its output power to avoid interference with a completely different system can be found in the original specification of the 802.11a standard. Here, the term 'transmit power control' (TPC) is used and it enables networks to lower the aggregate transmit power by 3 dB from the maximum regulatory limit to protect Earth exploration satellite systems (ESSS) operations [280].

9.5

Technical Barrier and Enabling Technologies

It is widely accepted that software radio is the key enabling technology for cognitive radio. Although any sensible definition does not require a cognitive radio to be implemented as an SDR for practical reasons, the extensive use of SDR technology will be required. The key reason for this is the sheer complexity of the systems that are being proposed or envisaged.

If SDR is an enabling technology for cognitive radios, then all the enabling technologies for SDR apply to cognitive radio. Many of these technologies are being studied for submission to Ofcom with this report. The key technologies are antennas and antenna processing, waveforms, MIMO techniques and signal conversion technologies.

As well as the technological hurdles that are common to SDR, cognitive radio has some of its own. Some of these obstacles are discussed in more detail in the following sections.

9.5.1 Radio Etiquette

The term ‘radio etiquette’ is used to describe the protocols and techniques that different devices will use to share the spectrum. If careful consideration is not given to this problem then the network would quickly become unusable [283].

For example, if a cognitive radio in the process of making a transmission detected an elevated interference level, the obvious (but selfish) thing to do would be to increase its transmit power. Unfortunately, if the other devices implement the same algorithm, they would respond by increasing their transmit power and an ‘arms race’ would ensue. In this situation, paradoxically, the best thing to do may be to reduce the transmit power. However, this would only work if all radios in the area used the same approach.

Another example for coexistence by using a measure of etiquette is collision detection and avoidance. If two radios transmit at the same time then there may be a ‘collision’, which makes it impossible for the receivers to decode the data. If the radios detect that a collision has occurred they may attempt to retransmit their data. However, if both radios detect a collision and immediately retransmit, the result will be another collision. The technique that is used to resolve this problem is called retransmit with random back off. Basically, each radio will wait a random amount of time (with a predefined maximum) before retransmitting. This relies on the fact that the probability that each radio will attempt to transmit at the same instant a second time is relatively low. Note that once one radio ‘gets in first’, the second radio would then wait until the first had finished transmitting, avoiding another unnecessary collision.

The example above is very simplistic compared to some of the negotiation problems that would have to be resolved to develop an effective system. Some researchers believe that intelligence will have to be built into the device so that they are capable of machine learning and self adaptation. It has been proposed that a field known as ‘game theory’ could be used to advance this area of cognitive radio.

9.5.2 Legacy User Protection and Coexistence

For obvious practical reasons, cognitive radios will not replace existing radios all at once. For this reason cognitive radios will have to coexist (and give priority to) legacy radio devices.

In simple terms, protecting legacy users means not transmitting when they are. However, the commonsense rule of “don’t transmit if you can decode” is inadequate; both from the standpoint of protecting legacy users and maximizing the potential for spectrum reuse [284]. The reason for this is that even though you cannot decode a signal from a legacy piece of equipment, it still might be present and interference may result if a new transmission is made.

Reference [284] also suggests that the best results that can be achieved are from a simple energy detector. It may not be possible to tell who is transmitting or what is being transmitted. However, it should be possible to detect that a transmission is being made and therefore to avoid interfering. Reference [284] also describes how, if the legacy user transmits a pilot tone or ‘beacon’ signal, legacy user detection may be greatly simplified.

UNCLASSIFIED

The key point is if cognitive radio systems have to coexist with legacy users, effective detection is crucial before the radio is allowed to operate.

9.5.3 Spectrum Trading

Although it has been shown that DSA would undoubtedly offer benefits in terms of spectrum utilisation and effective throughput, careful consideration has to be given to the commercial reality of such a scheme. Exactly how this trading would be done, and on what basis the costs would be calculated needs to be thought through. For example, questions such as “will a spectrum broker be required and, if so, who will be that broker?” need to be answered before spectrum trading can begin.

It may prove difficult to build a mechanism for spectrum trading that limits anticompetitive trading. For example, one service provider might refuse to sell unused spectrum to another provider for commercial reasons. Anticompetitive behaviour is not in the interests of consumers, so steps need to be taken to prevent it from occurring.

The mechanisms for spectrum trading will become increasingly difficult as the scheme that is adopted becomes increasingly dynamic.

9.5.4 Regulatory Issues

Ofcom serves to further the interests of citizen-consumers as the communications industries enter the digital age [285]. Increased spectrum utilisation and the resulting increase in the number of wireless services available is undoubtedly in the interest of the consumer. Although cognitive radio is becoming increasingly viable in terms of technology, the current regulatory position would significantly impede its development and deployment. Whilst there are very good reasons for the current regulatory framework, the current regulatory position has to be viewed as a hurdle to the successful deployment of cognitive radio and the benefits it might bring.

Although cognitive radio devices and DSA offer significant advantages to radio users, there are some potentially serious implications. The current situation of static allocations means that it is easy to determine who should be transmitting and at what maximum power. Therefore it is relatively easy to investigate complaints and hence effectively police the spectrum. However, if cognitive devices are introduced then situations might arise where several different types of devices or services are attempting to access the same spectrum, at the same location and at the same time. Determining who has priority might be no mean feat. Indeed, simply determining exactly who is transmitting might be a challenge.

Providing that sufficient radio etiquette is maintained and systems are designed properly, these systems should all be able to coexist. However, procedures will have to be put in place to deal with devices that are either intentionally or unintentionally operating incorrectly. This will require a significant amount of effort from organisations such as the FCC in the US and Ofcom in the UK.

UNCLASSIFIED

On the other side of the argument, protection will need to be put in place to protect cognitive radio devices from legacy devices. The IEEE is concerned that legacy services may intentionally disrupt cognitive devices [286]. They suggest that existing users should be prevented from transmitting just to stop cognitive radios from operating or forcing them to move to a different band.

9.6

Conclusions

In this chapter we have shown that the current approach to radio spectrum allocation is becoming increasing inadequate. Much of the allocated spectrum is underused while there is very little ‘useful’ spectrum left for future applications.

Promoting the concepts of spectrum trading and liberalisation and opening up the spectrum to market forces should enable significant gains in spectral efficiency to be realised. Different proposals for the ways in which this could be done have been discussed. The simplest idea was to change the current regulations to allowing the primary license holder to sell or lease their spectrum to other users. We note that currently a primary license holder is not permitted use spectrum allocated to them for any purpose other than that for which it was originally licensed.

The next step from long-term spectrum trading would be towards a more dynamic method of spectrum sharing. Research has showed that this would bring significant advantages. However, it might require fairly complicated technology to enable it to be implemented. In order to implement the most dynamic visions of spectrum trading it is likely that cognitive radios will be required.

Several definitions of a ‘cognitive radio’ exist, and most seem to restrict them to making dynamic use of the radio spectrum. However, it was shown that even with a fixed frequency allocation, there is plenty of scope for cognitive radios to offer improvements over conventional radios. Although a fully cognitive radio has yet to be deployed, there are several existing systems that do have some awareness of their environment and can adapt accordingly.

Finally, some of the technological hurdles that will have to be overcome before CRs can become reality have been discussed. CRs will most likely be based on SDRs. Therefore, any technological hurdle for SDRs is also a hurdle for CRs. Other obstacles that may hinder the development of CRs are the development of techniques for the successful sharing of spectrum (known as radio etiquette), techniques for protecting legacy users from interference, and the actual mechanisms for trading radio spectrum. It has been noted that CR and dynamic spectrum trading are not possible within the current regulatory environment.

10 Regulatory Issues



By Tim James, Multiple Access Communications Ltd.

10.1 Introduction

This chapter considers regulatory issues in the context of software defined radio (SDR). SDR represents a significant milestone in the evolution of radio equipment. It moves away from 'traditional' radio architectures with relatively inflexible, dedicated signal paths toward more generic, undedicated radio platforms with software to define and, if necessary, redefine the required functionality.

The potential advantages offered by SDR are many. In particular, an SDR has the ability to be reconfigured quickly and efficiently at the time of use rather than at the time of design and/or manufacture. The use of standard, generic SDR platforms will also facilitate the rapid development of new, innovative radio systems at much reduced cost. Furthermore, reconfigurable SDRs are more 'future-proof' than radios with more traditional architectures. Shifting more and more signal path functions from the analogue to the digital domain implies more deterministic radio characteristics. This should mitigate some of the risks that might otherwise be associated with deploying new radio configuration updates on different equipment without performing rigorous manufacturing tests.

SDR is not without its drawbacks. Programmable logic is typically larger, costlier and has increased power consumption compared to optimised, dedicated hardware. Such constraints are likely to prevent the widespread use of SDR in user equipment for the foreseeable future. As with virtually any software there is the potential for bugs to cause havoc in prematurely released and/or improperly tested software. Considering the reprogramming of SDRs in the field, there are significant question marks over how this should be implemented and how downloads might be authorised to prevent the use of non-conforming, non-type-approved and/or malicious software configurations. For example, considering an extreme case, an SDR might easily be reconfigured to operate as a jammer. There are also, of course, significant regulatory issues concerning how best to regulate the use of SDRs and how to manage the type approval process.

This section considers these regulatory issues, summarises the FCC's reaction to SDR and discusses the impact of SDR on the European Union's (EU's) radio equipment and telecommunications terminal equipment (R&TTE) directive.

There is a separate study, commissioned by Ofcom and undertaken by the University of Surrey in conjunction with Fujitsu, which discusses Regulatory issues in much greater detail [287].

10.2 What is SDR and at What Point Does the SDR Part Stop?

In order to discuss possible regulatory issues associated with SDR it is first necessary to define exactly what SDR is. This has been discussed in previous sections but is summarised again here for convenience.

UNCLASSIFIED

The exact definition of SDR would appear to be somewhat ambiguous. Does the ‘software defined’ part describe how the radio is implemented or how the radio is controlled? Further, is it a requirement of SDR that it can be field programmable, either over the air or via some other means? Finally, to what section of the signal path does SDR apply?

Two ‘visions’ of SDR exist. One vision of SDR is at a high level and presents the concept of a black box that can be reconfigured rapidly, on the fly using internally held configuration data or, ultimately, using configuration data downloaded over the air. Thus, the focus here is on reconfigurability, perhaps leading to cognitive radio applications. Under this definition, SDR can include radios with a ‘traditional’ architecture but incorporating a highly flexible RF front-end and multiple baseband processing modules, all controlled and configured by software. Indeed, this definition can already be applied to many modern cellular handsets. Another vision of SDR is at a lower level and presents the concept of a radio system in which part or all of the signal path functionality is implemented digitally using ‘generic’ signal processing entities such as FPGAs and/or DSPs. Thus, the focus here is on the method of implementation. Reconfigurability is not a requirement (although, essentially, functional changes can be made without modifying the hardware). These two ‘concepts’ of SDR are not necessarily compatible. Although both concepts might be considered valid, the first has more of a commercial spin whilst the second has more of an engineering spin and is perhaps closer to the original vision of ‘true’ SDR where the signal path is implemented in the digital domain right up to the antenna. Note that the key to all SDR is that the characteristics and functionality of the radio are defined by software and can be altered without any changes to the hardware.

Which elements of a radio does SDR relate to? Consider the seven layer open standards interface (OSI) reference model, shown in Figure 10-1. The communications engineer is normally interested in the lower three layers. The physical layer, which is concerned with the transmission of raw data bits over the ‘physical’ communications channel and, in a radio, defines aspects such as the radio frequency, modulation scheme, symbol rate, etc. The data link layer, which sits above the physical layer and handles packet framing, multiple access and data duplex to present a transparent ‘bit-pipe’ to the processes implemented within the higher layers. And the network layer, which is responsible for controlling the setting up, management and termination of data channels as required. So, which parts of the OSI reference model are applicable to SDR?

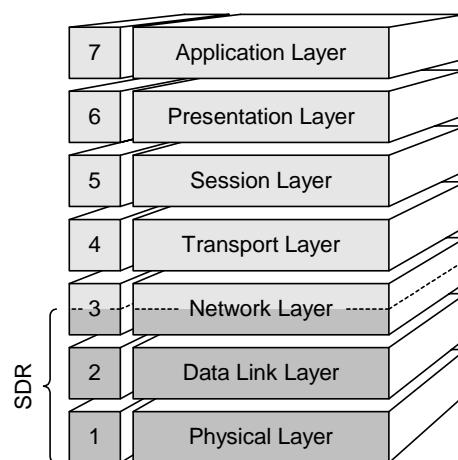


Figure 10-1: The seven-layer OSI reference model

It is suggested that SDR relates to functionality within layers one and two of the reference model and, in some cases, parts of layer three. All higher level functionality might well be implemented in software and may even be reprogrammable, but this is not SDR. Certainly SDR has no role at the application layer. Giving the user the ability to download new software onto his/her handset, perhaps to implement some location based service, does *not* make the handset an SDR.

So how does this translate to hardware? A SDR is, ideally, a generic, i.e., non-technology-specific, hardware platform whose physical layer function is defined in whole or (more practically) in part by software. Note that software in this instance encompasses firmware, e.g., FPGA configuration data. A simplified block diagram of an SDR platform, realisable using today's technology, is shown in Figure 10-2.

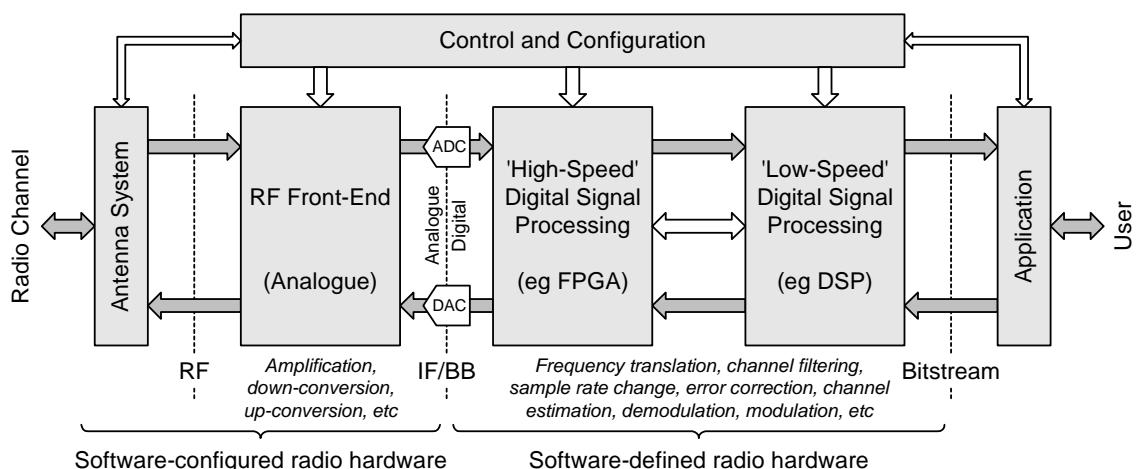


Figure 10-2: Simplified block diagram of a realisable software defined radio platform using today's technology

As with all radio systems, some form of antenna system (passive or active) is needed to handle the electrical/air interface. Current limits in technology mean that most SDR platforms will, for the foreseeable future, incorporate a 'traditional' RF front-end consisting of LNAs and downconverters on the receive side, and upconverters and power amplifiers on the transmit side. The analogue and digital domains are bridged using high-speed ADCs and DACs operating at IFs or at baseband (BB). Note that, whilst not truly software defined, the RF front-end may still be software *configured* to enable operation over multiple RF bands and at different transmit powers. Such software configured hardware is thus still in keeping with the concept of SDR.

Using the technologies currently available, an SDR will receive and/or transmit a digitised IF or baseband waveform. 'High-speed' digital signal processing will typically be required at the digital/analogue interface and might be implemented using FPGAs or ASICs with parameterised signal paths, for example. As the sample rate is reduced, DSPs might be used to handle 'low-speed' digital signal processing culminating in the raw user bit stream used by higher level processes.

UNCLASSIFIED

Initially, configuration data are likely to be target specific. For example, a radio configuration might consist of a configuration bit file for a particular FPGA device and a binary executable for a DSP. Moreover, the configuration data incorporates platform-specific data such as the FPGA pinout and DSP memory map. Ultimately, however, new methods might be developed to describe the signal path at a high level which would enable SDR to be defined in a more generic manner. For example, rather than synthesising a finite impulse response (FIR) channel filter to be implemented within a particular FPGA, the configuration data might simply describe an N -tap FIR filter with a set of filter coefficients. Once downloaded, the SDR platform would then implement the specified filter using ‘standard’ components from a built-in signal path library. Such an approach would offer two significant benefits. First, radio configurations stop being platform specific, which in itself would bring numerous advantages. Second, by parameterising the signal path, the size of each radio configuration might be reduced, which would greatly reduce the download time for new radio configurations.

As technology advances, the functionality of the RF front-end will be increasingly emulated directly in the digital domain, moving nearer and nearer towards the ‘true’ SDR.

10.3 What are the Benefits of SDR?

Why is SDR so important? What benefit does it give to the manufacturer, regulator, network operator and user? We consider the benefit to each part in the following sections.

10.3.1 The Manufacturer

From the point of view of the equipment manufacturer, SDR might facilitate the streamlining of a large portfolio of dedicated radio hardware products, targeted at a multitude of radio technologies, down to a more condensed range of generic radio platforms. Thus, in the long-term, the manufacturer might realise significantly reduced hardware development costs, i.e., the cost of developing a single programmable radio platform can be spread over many end products. Not only are development costs spread, the number of hardware product lines are condensed, allowing the streamlining of the manufacturing process.

Having developed a common, generic programmable radio platform, the manufacturer can then target a wide range of radio technologies simply by selecting the appropriate configuration software, either developed in-house or subcontracted out to third party developers. Furthermore, with only software development required, the development of new products supporting new/variant radio technologies can be accelerated greatly. Similarly, the time between additional features being defined in new releases of a particular radio technology specification and their being integrated with the product line might be reduced. Not only might these advantages give a manufacturer a head start on its competitors, but shorter development times imply lower development costs.

UNCLASSIFIED

A further benefit of SDR to the manufacturer lies in the ability to adapt development cycles more commonly associated with the development of application software. When developing new radio equipment and in particular radio equipment designed to support emerging radio technologies, the initial software release need only support a minimal function set. Further software development can then continue in parallel with hardware testing and initial trials with software updates incorporating additional radio functions being made as and when they become available.

Whilst the programmable devices needed to realise generic radio platforms might be more expensive than today's application-specific chipsets, a single programmable device might be used to replace the multiple baseband processors found in modern multi-mode radios. SDR might also create a market for third-party software developers as it would create a range of original equipment manufacturer (OEM) SDR modules upon which more bespoke systems could be developed.

10.3.2 The Regulator

Regulators such as Ofcom in the UK and the FCC in the US are keen to develop the concepts of spectrum liberalisation and spectrum trading to foster the efficient and innovative use of the available radio spectrum. Taking these concepts a stage further is the concept of CR, that is, radios that incorporate limited intelligence to adapt to the quality of service requirements of the user and the local radio environment. Thus, a CR might be able to dynamically switch modulation scheme and both frequency and time domain behaviour to exploit 'gaps' in congested spectrum and/or find unused spectrum in which to operate. Hence CR has the potential to truly revolutionise the efficient utilisation of the radio spectrum for telecommunication applications and is, therefore, of great interest to the regulator.

SDR is not a prerequisite to spectrum liberalisation, spectrum trading or CR. However, the flexibility of an SDR may well mean that SDR becomes an enabling technology that makes such concepts a widespread reality. Consequently, the regulator needs to be receptive to SDR and, if necessary, modify current regulatory procedures to foster their development and deployment if it wishes to accelerate the acceptance and adoption of spectrum liberalisation and trading.

10.3.3 The Network Operator

From the point of view of the network operator, SDR has the potential to realise 'future-proof', multi-mode base station radios and user equipment. Thus, an operator might upgrade its base station radio equipment and/or the equipment of its users to incorporate the latest incremental changes to a radio technology standard. Moreover, the operator might be able to switch to a 'next-generation' technology both rapidly and without having to replace existing base station equipment. Furthermore, by enabling the remote reconfiguration of base station equipment, the operator can maintain and update its network infrastructure without having to gain physical access to its equipment. This might be especially important when considering indoor picocell base stations that are often inaccessible after they have been commissioned. Notwithstanding, it is acknowledged that even a reprogrammable SDR can only ever be 'future-proof' to a limited extent; as radio standards become more computationally intensive, the capabilities of 'older' SDR platforms will be exceeded, requiring the inevitable hardware upgrade.

UNCLASSIFIED

By shifting baseband signal processing from dedicated hardware resources into generic signal processing resources, base station system complexity and size might be reduced. Thus, SDR might ultimately lead to lower equipment costs and improved system maintainability and reliability. Again, this might be important when considered in the context of indoor picocell base station equipment.

Finally, by investing in SDRs, network operators will be able to position themselves favourably for when spectrum liberalisation, spectrum trading and CR become reality by equipping themselves with the flexible network infrastructure required for the implementation of such concepts.

10.3.4 The User

In the long-term, the user might observe many benefits of SDR. Any cost savings passed from the manufacturer to the network operator should ultimately be passed to the user; although, given the high level of handset subsidy made by the UK cellular operators, it is acknowledged that this may simply result in a reduction of subsidy levels. By enabling the network operator to deploy software updates following amendments to the radio technology standard, the user might gain access to the latest service innovations sooner. Finally, if SDRs do indeed lead to improved system reliability and maintainability, the user should receive a better quality of service from the network; perceived quality of service and customer satisfaction is most certainly a key factor of brand loyalty.

Initially, cost and power consumption will probably prevent the widespread availability of SDR user equipment. Ultimately, however, market forces and innovative radio technologies might lead to the development of reconfigurable handsets. Once this takes place, the user would be less restricted when selecting a network to use when roaming outside the UK; furnished with the relevant software upgrade, the user would be able to access virtually any network worldwide (providing the relevant inter-operator roaming agreements were in place). Within the UK, a reconfigurable handset might dynamically switch mode to provide the user with an optimal compromise between quality of service and cost of service, especially when considering data services.

10.4 Why is SDR Such a Potential Regulatory Minefield?

So, we know what SDR is and why it is so important. Why does it pose such an issue to the regulator? A SDR can, theoretically, be reconfigured 'on the fly'. Furthermore, the radio's function is not necessarily fixed at the time of manufacture; new configurations might be downloaded either from the web, using removable data storage media or over the air. Thus, the traditional approach to type approving new radio equipment before it is made commercially available is no longer a realistic approach or even necessarily possible. Therefore, either restrictions need to be imposed on the application of SDR or the regulatory aspects need to be reviewed. Given the potential importance and benefits offered by SDR, the first option would block innovation and might prevent the development and acceptance of concepts such as spectrum liberalisation, spectrum trading and CR in telecommunications. Therefore, the regulatory aspects need to be reviewed.

UNCLASSIFIED

The potential enormity of the issue of regulating SDR becomes apparent when considering the number of ‘dimensions’ in which the function of a software defined radio transceiver might be reconfigured. These include:

- Modulation scheme, e.g., phase-shift keying, frequency-shift keying, Gaussian minimum-shift keying, etc.
- Channel spacing and raster
- User data rate(s) and channel bandwidth
- Frame and packet timing, structure and format
- Logical and physical channel definition
- Duplex operation, i.e., frequency-division duplex or time-division duplex
- Multiple access methodology, e.g., TDMA, FDMA, CDMA, OFDMA, etc.
- Frequency band of operation
- Out-of-band emission control, i.e., pulse-shaping filters and transmit power spectral density masks
- Permitted transmit power levels
- Power control mechanisms
- Channel estimation and equalisation techniques
- Channel coding and forward error correction techniques.

Note that this is not intended to be an exhaustive list. All of these dimensions may be reconfigured in a true SDR and all, to a greater or lesser extent, have the potential to impact how effectively the available spectrum is utilised, how reliably different transceivers can communicate and how much interference is caused to other users. As an example, out-of-band emission control will have a direct impact on the interference caused to users operating on adjacent channels. Inefficiently conceived and/or implemented power control, channel equalisation and/or channel coding mechanisms might lead to higher transmit powers which may, ultimately, have a detrimental effect on the capacity of the host network and spectral efficiency as a whole; although this is more an issue for the operator rather than the regulator.

10.5 Regulatory Aspects for Different SDR Applications

Before the various regulatory aspects can be considered it is first necessary to consider how, initially at least, SDRs might be deployed in practice. The ultimate vision of SDR might be undedicated user terminals, operating in an ad-hoc manner, intelligently reconfiguring themselves to adapt to the current radio environment and user requirements. Here, the terminals will download new implementation configurations from SDR software servers or even use spectrum analysis techniques to automatically adapt to the received signal, and will only be loosely constrained as to the frequencies on which they operate. This is really the long-term vision of cognitive radio; here SDR is more of an enabling technology rather than the end product. In the short term, however, the reality will be a bit easier to comprehend.

UNCLASSIFIED

Let us consider realistic application scenarios and the associated regulatory aspects. It is necessary to consider SDRs in the context of both client/server (e.g., cellular) networks and, thinking ahead, ad-hoc, peer-to-peer networks. It is also worth noting that, although user terminal SDRs should be considered, the reality is that cost and power constraints are, initially at least, likely to restrict SDRs to base station applications.

10.5.1 Base Station SDRs

For the foreseeable future, the most likely application of SDR will be in the implementation of base station equipment. These have the potential to offer cost-effective, ‘future-proof’ equipment to the network operator. Other benefits include reduced system complexity, which will help improve system maintainability and reliability.

As stated above, a key advantage of a base station SDR is its potential as a future-proof solution. Thus, it can be updated with bug fixes, changes resulting from amendments to the radio technology standards and, ultimately, complete retargeting to different radio technologies. It is reasonable, therefore, to assume that base station SDRs might be updated with new software after deployment. In this case, how should such equipment receive type approval?

Realistically, any hardware/software configuration change that can potentially affect either the transmitted signal or receiver implementation needs to be retested before it can be deployed in the field. Who is responsible for getting type approval for new hardware/software configurations? The equipment vendor? The software vendor? Or the operator? Furthermore, who is interested in what? This question is even more pertinent with the advent of spectrum liberalisation.

With the regulator currently allocating spectrum to a particular operator for a particular purpose, the regulator would need assurance that equipment performs as intended before updates could be deployed. However, with the advent of spectrum liberalisation the goalposts are moved. The regulator is now less concerned with what the operator does with their allocated spectrum provided that the interference caused to other spectrum users remains below predefined limits.

Thus, as before, the regulator requires assurance that a software update will not result in an increase in out-of-band interference. However, whether or not a software update introduces incompatibility with the other equipment operating on the operator’s network becomes primarily an issue for the operator to resolve.

Note that fully software defined base station radios are in existence today. PicoChip Designs Limited have demonstrated fully functioned 3G base station reference designs on its picoArray devices [288]. In the US, Vanu Inc. has been the first company to obtain type approval of an SDR base station through the FCC [289].

UNCLASSIFIED

10.5.2 User Equipment SDRs with Preloaded Configuration Data

Moving to consider user equipment SDRs, the simplest category will be that of generic radio platforms that are programmed at manufacture to implement the required functionality. Thus, once off the production line, the radio's functionality cannot be changed (at least, not by the user). Note that this is not to say that such a radio might not be multi-modal and/or multi-band and cannot be reconfigured on the fly. Such equipment is analogous to 'conventional' radio equipment, differing only in the method of implementation. Therefore, there is no reason why the approval process of such equipment need be changed from that which is used currently.

10.5.3 Reconfigurable User Equipment SDRs for Use in Client/Server Networks

In the future, when reconfigurable SDR user equipment becomes more viable, network operators might start to see SDRs operating on their networks. In theory SDRs are good because they can be retargeted to operate on virtually any network. However, there are many technological and regulatory aspects to consider.

Unlike SDR base station equipment, over which the operator has complete control, there might be very little that an operator can do to manage the use of SDR user equipment on its network. Whilst new technology standards might be developed specifically to support the operation and authorization of SDRs, there is no such support in existing standards, e.g., a base station would not be able to distinguish a conventional GSM phone from an SDR handset emulating the functionality of a GSM handset.

There are perhaps two primary methods of reconfiguring an SDR as shown in Figure 10-3. The first is to download new configuration data over the air from a central SDR software server. The second is to reprogram the device 'off line', perhaps via a download from the Internet or through the insertion of a high-density data card, e.g., a SD card or a 'mega' SIM card. There are, however, significant pros and cons relating to each method.

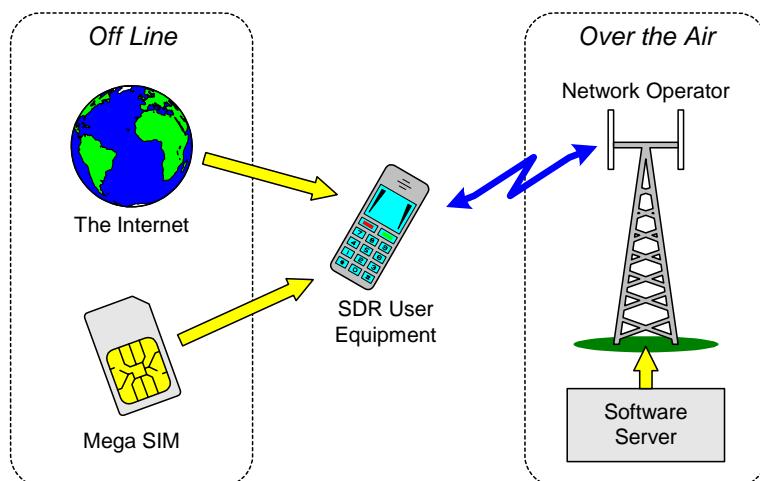


Figure 10-3: Updating software on SDR user equipment

UNCLASSIFIED

First let us consider over-the-air downloads from a central SDR software server. Such a server would be accessed through the network operator. This provides a number of advantages. From the point of view of the network operator it is easier to control what configurations of hardware and software are permitted to operate on its network. From the point of view of the ‘typical’ user, this should greatly simplify the task of reconfiguring an SDR handset. Ultimately, it is quite possible that the user need not even be aware of software updates; the availability of software updates might be detected and downloaded automatically. A significant drawback of this approach, however, is that a ‘chicken-and-egg’ situation is created when the user roams to another network. Without a compatible software configuration the user will not be able to access the software server to download the appropriate radio configuration data. Thus this potentially would require a globally harmonised radio interface through which an SDR could search for and request software updates. Such an approach would also require the operator to implement and manage the software server, the cost of which might be significant.

The alternative to over-the-air software updates is to allow the user to reprogram his/her handset ‘off line’. This is undoubtedly a simpler method as it avoids the need for a harmonised interface to be defined in order to access on line software servers around the globe. This means that a user could reprogram his/her handset to operate on non-native networks once in place without needing some internationally realised software server network. However there are some significant drawbacks. First, the reprogramming method is less user friendly to the ‘typical’ user. Perhaps more importantly, however, is that the network operator now loses any method of managing SDRs on its network. In particular, there is greater scope for non-type approved, ‘bootleg’ software finding a market (particularly with Internet downloads enabled).

Over-the-air downloads are likely to be a distant reality due to the greatly increased complexity (when compared to off line updates) and the need for a recognised (and implemented) software server ‘standard’. Ultimately it is probable that a combination of both methods will be required; over the air updates for everyday operation with an alternative means of reprogramming for instances where an ‘on line’ software server cannot be accessed. The development of a means to deliver over the air updates might ultimately be driven by the emergence of CR networks.

Before over-the-air software downloads can become commonplace, there is a requirement for a standardised method of authorising the use of software on SDR hardware. This might entail some form of encryption mechanism whereby an unlocking key is released by a central regulatory body following the registration of new software. The issue with this approach, however, is that it imposes certain implementation requirements on the software and hardware vendors, i.e., all equipment will need to incorporate common operating system functionality to implement the locking mechanism.

Furthermore, a central authorisation repository might be required. Note that there is significant interest in this topic. Examples include the work by Michael et al [290], Brawerman et al [291] and the SDR Forum [292].

UNCLASSIFIED

10.5.4 Reconfigurable User Equipment SDRs for Use in Ad-Hoc, Peer-to-Peer Networks

With ad-hoc networks, in which users communicate at a peer level, the role of the network ‘operator’ may be much diminished. Therefore, whereas in a client/server network the network operator might take on some of the responsibility for ensuring only verified hardware/software combinations are operated, in an ad-hoc network it may be much more difficult to regulate the use of software radio.

Again, therefore, there might be a requirement for a standardised method of authorising the use of software on SDR hardware. Users accessing a software database would have to download the configuration data and then obtain an authorisation key before the hardware will reconfigure itself with the new data.

10.5.5 Reconfigurable User Equipment SDRs for Use in Cognitive Radio Networks

The concept of cognitive radio covers a wide range of operating scenarios. Perhaps the simplest form of CR describes radios that use techniques such as frequency hopping, power control and/or dynamic burst timing to implement interference avoidance and/or minimise interference caused to others. Note that the power control mechanisms implemented in CDMA networks represents a basic form of CR as the radio dynamically adjusts its power to achieve the requested quality of service whilst keeping interference to other users to a minimum. In such applications, CR is part of the implemented radio technology and is not an issue specific to the use of SDR.

In the future, CR might be implemented in client/server networks supporting SDR user equipment at the base station. Here, under the control of the base station, traffic channels would be set up in any available spectrum, perhaps leased from some third party, using a radio technology optimised for the current environment and requested service type, e.g., speed, video, data, etc. In using the base station to provide the intelligence required for CR, the network operator retains full control over transmissions in ‘other’ bands. The only user equipment transmissions that are not strictly controlled by the base station are random access attempts which would only be allowed on predefined, licensed channels. The base station might allocate traffic channels in shared/reusable spectrum by interrogating a location-addressed national database and/or signal measurements provided by user equipment. Using the downlink as a ‘beacon’ signal, a failsafe mechanism is implemented that would prevent user equipment transmitting illegally. This scenario is made possible by permitting spectrum trading and by liberalising the use of spectrum. Here, SDR is an enabling technology. Significant development of this concept is required before it can become reality. However, considering the SDR element, there should be no SDR-specific regulatory issues that do not exist when considering using over the air reconfigurable SDR in non-CR networks.

UNCLASSIFIED

In ad-hoc, peer-to-peer networks user equipment might use CR techniques to automatically select an appropriate radio configuration, either from a locally stored set or downloaded automatically from a software server. Controlling spectrum access and authorising the various software configurations would present significant regulatory issues. Again, however, these issues are not significantly different from those presented by SDR in non-CR ad-hoc networks. The only added dimension is that of constraining which configurations can be used in which areas at any particular time. In particular, if such networks are to be allowed to reuse public safety spectrum, for example, a means to managing access that provided a failsafe means of clearing the spectrum in case of an emergency would be required. This topic is not specific to SDR and is outside the scope of this study.

Ultimately, a CR might automatically and dynamically adapt its operation using built-in signal path building blocks, using spectrum analysis techniques to dynamically adapt to different radio technologies. Thus a CR terminal within range of a radio operating some unknown radio technology will analyse the received signal and, using a library of standard signal path building blocks, automatically configure itself to emulate a compatible transceiver. The realisation of such a vision is certainly a very long way off and may never happen. First, SDR has to mature and low power handheld equipment with significant processing power has to become a reality. CR implemented using radio software from a configuration database will be an intermediate step. Needless to say, however, SDR with the ability to dynamically alter its configuration would be virtually impossible to regulate.

10.6 Type Approving SDR Applications

The main regulatory headache concerning SDR is that of equipment type approval. Unlike conventional radio equipment whose function is typically defined at the time of manufacture, the function of SDR is not necessarily defined until after manufacture and can potentially be changed at will thereafter. Thus, the traditional process of acquiring type approval for a new piece of radio equipment before making it commercially available is not necessarily suitable when considering SDR.

This issue is potentially compounded when it is realised that it would be difficult, if not impossible, to verify the continued type conformance of SDRs reprogrammed in the field. This is especially the case if the software update retargets the hardware at a completely different radio technology to that for which it was originally intended. In other words, whilst it might be possible to test an SDR platform at the time of manufacture using a range of software test configurations, it will not be practical to exhaustively test every unit for all potential radio configurations. It is acknowledged that the deterministic behaviour of digitally-implemented radios represent a significant advantage over traditional analogue hardware because variation between units can, theoretically, be eliminated. However, for the foreseeable future, all SDRs will still require ‘conventional’ RF front-ends with their associated nonlinearities and unit-to-unit variations. Therefore, type approval for radio configurations released as field updates will have to be acquired purely on the basis of test data generated from a batch of ‘typical’ hardware examples. Production line testing may no longer be able to provide a safety net for non-conforming radios.

UNCLASSIFIED

One possible solution might be to require the inclusion of built-in self test (BIST) functionality, where a proportion of the transmitted power can be looped back on the receive path for analysis; detection of non-conformance might then ‘block’ the further use of that particular radio configuration. However, such an approach is likely to add significantly to the complexity and hence cost of the system. It might, therefore, be that the threat of the occasional nonconforming radio is outweighed by the advantages to be gained by allowing the use of reconfigurable radio platforms.

The scope of these type conformance issues is increased if it is considered that one manufacturer might develop and market the hardware whilst a third party develops the radio software. Who is responsible for guaranteeing regulatory conformance? The hardware vendor or the software vendor?

On its own, i.e., without software, an SDR platform has no function and, with the exception of satisfying basic electromagnetic compatibility (EMC) requirements, cannot undergo type approval as there is nothing to prove it against. Therefore the hardware manufacturer can argue that they are not responsible for getting type approval.

A software developer might be able to prove through simulation that his/her radio implementation *theoretically* meets the minimum stated radio performance. However, the hardware will almost certainly cause some degradation from the theoretical performance, e.g., through the introduction of local oscillator phase noise with the wanted signal during mixing or through nonlinearity in the signal path. However, requiring the software developer to obtain type approval for all combinations of new software with existing hardware versions might discourage would be developers, especially if his/her business case requires him/her to target multiple hardware platforms.

Is there an alternative that could relax the requirement for independent software developers to test new software configurations on representative hardware samples? If a relatively simplistic yet complete model of imperfections in the radio hardware could be devised, might the hardware manufacturer have a model against which it can type approve the radio characteristics of its product? Such a model might incorporate manufacturing tolerance data to enable the probability of non-conforming radio equipment to be determined. The software vendor might then be able to rely on simulations incorporating this hardware model to gain type approval for new software releases. Significant further work and study would be required in order to realise such an approach and considerable effort would be needed to validate such a methodology. Long-term, however, the development of such an approach might be beneficial.

Note that, initially, configuration data are likely to be platform specific. Therefore to require the testing of new software on all compatible hardware is, perhaps, not too onerous a task. Long-term, however, the industry is likely to pursue the concept of an SDR platform with a built-in library of parameterised signal processing modules, together with a high-level description language that allows common configuration data to be used to target multiple SDR platforms. Thus the software configuration describes the signal path in an unambiguous manner at a high level that can then be assimilated by the platform using predefined building blocks. This might be a far-off vision but, were it to become reality, the task of verifying software updates on all compatible hardware platforms would soon become impractical and alternative means would be necessary.

10.7**Equipment Type Approval in the United States by the FCC**

The FCC in the US has identified the potential limitations that its current method of regulating telecommunications equipment might impose on SDR. To this end, the FCC has undertaken extensive studies into SDR and proposed new type approval procedures as a result.

Before proposing any SDR-related rule changes, the FCC first had to define a software defined radio. The (current) FCC definition of a software defined radio, as defined in FCC 01 264 [293], is as follows:

“A radio that includes a transmitter in which the operating parameters of frequency range, modulation type or maximum output power (either radiated or conducted) can be altered by making a change in software without making any changes to hardware components that affect the radio frequency emissions.”

Note that this definition focuses on the transmitter rather than the receiver. Initially this makes sense as it is the transmitter as the radiating device that ultimately will cause any interference to other users. From a type approval standpoint, therefore, this is probably a satisfactory definition for SDR. It should be noted, however, that receiver performance in modern interference-limited radio systems employing closed-loop power control has the potential to affect overall network capacity. Consider a software defined implementation of a CDMA receiver utilising closed-loop power control. If changes to the receiver implementation degrade its sensitivity, the receiver will request the output power of the originating transmitter to be increased. This will increase interference to other users and will reduce cell capacity and/or coverage as a result. Thus, although changes to a receiver implementation cannot cause interference to other users directly, such changes could theoretically have secondary (albeit relatively minor) effects. It is acknowledged, however, that this is not necessarily an issue for the regulator. The regulator requires assurance that *maximum* transmit powers are not exceeded. Capacity degradation caused by poor receiver performance is more of a concern to the network operator which might see a loss of revenue resulting from inefficient utilisation of its spectrum.

Receive-only radio equipment does not normally require type approval; as, by its nature, it cannot cause significant interference to other radio users. However, SDR is by no means restricted to transmitters and so it is perhaps strange to exclude software defined receive-only radio equipment from any definition of SDR. A subtle rewording of the above definition might therefore be appropriate to incorporate software defined receiving equipment also.

Note that the FCC definition of SDR does *not* require that the radio is field re-programmable. Thus, under this definition, SDR includes equipment which uses software installed at manufacture to configure a generic radio platform.

Prior to the FCC's study into SDR, the FCC defined two classes of ‘permissive’ changes that could be made to equipment without requiring it to re-undergo full type approval. A third class of change has been defined following the FCC's SDR study. The three classes of permissive change as defined in Title 47 of the Code of Federal Regulations (47 CFR §2.1043 [294] are, in summary, as follows:

- Class I – Modifications which do not degrade the characteristics reported under the original type approval application

UNCLASSIFIED

- Class II – Modifications which do degrade the reported characteristics but still meet the minimum requirements of the applicable rules
- Class III – Software modifications to a software-defined transmitter that alter the previously approved frequency, modulation type or output power.

No filing is required for Class I permissive changes. Equipment incorporating Class II or Class III permissive changes may only be marketed following approval of documented changes submitted to the FCC. Once approved, Class III changes can be made to equipment incorporating Class I changes. However, Class III changes cannot be made in addition to Class II changes without reapplying for full type approval.

An important point regarding the FCC's decision regarding these permissive changes is that the FCC state that only the applicant for the original equipment authorisation may submit Class III changes. The justification for this is that the owner of the authorisation certificate is responsible for the conformance of the end product and it would, therefore, be unreasonable for the certificate owner to be responsible for software changes made by a third party. This is of course a sensible approach. However, it does effectively prevent third parties contributing software for use on generic radio platforms. Moreover, it restricts the user of an SDR from choosing configuration software not provided by the OEM.

This raises an interesting question. Suppose a particular OEM SDR radio platform was authorised by different third parties, Vendor A and Vendor B, with his/her own developed configuration software. What mechanism would prevent the owner of a radio branded by Vendor A from loading software developed by Vendor B? The FCC requires the applicant for equipment authorisation to demonstrate how his/her equipment prevents the downloading of unauthorised software. Note that, although groups including the SDR Forum and ETSI are reportedly developing encryption and authorisation mechanisms for SDR, the FCC leaves the implementation of any authorisation mechanism up to the manufacturer.

There are two other findings of the FCC's review of SDR that are significant. The FCC is to allow equipment to store and display the FCC authorisation number electronically, i.e., on an light-emitting diode (LED) or liquid-crystal display (LCD) display, rather than requiring physical marking of the equipment [295]. This is significant because this avoids the otherwise implied requirement that equipment is returned to an authorised body to install software updates and modify the FCC enumeration. An electronic display of the FCC number means that configurations requiring a new FCC authorisation number *can* be downloaded to equipment *after* the point of sale. Note that this method of electronic labelling presumably is the key that will permit a user to reconfigure his/her equipment with configuration data from different third parties. A final conclusion by the FCC is with regard to the conformance testing of SDR. The FCC have decreed that, for the time being at least, SDR is still too young a technology to reliably allow type approval to be determined from the software and hardware elements in isolation. Consequently, the FCC requires that conformance testing be conducted using all intended hardware/software combinations.

The FCC has certainly taken significant steps in recognising the importance of SDR and in reviewing its equipment authorisation procedures to distinguish hardware and software changes. However, it is questionable whether the new procedures are still too inflexible and will prevent third parties from exploiting new business opportunities in SDR.

10.8 Equipment Type Approval within the EU and the R&TTE Directive

Within the member states of the EU, the type approval of radio equipment is currently governed by the R&TTE Directive 99/5/EC [296]. Member states had until April 2000 to implement the directive, which replaced previous national type approval systems.

Under the R&TTE directive, the manufacturer is responsible for ensuring type conformance. Two classes of equipment are defined. Class 1 equipment operates harmonised standards such as GSM, TETRA, DECT, etc. in harmonised frequency bands and can be marketed without restriction. Class 2 equipment operates non-harmonised standards or operates in non-harmonised frequency bands and may be subject to restrictions in certain member states. All approved equipment must carry the 'CE' mark as a sign of conformance. In addition, Class 2 equipment must carry the 'alert' symbol, shown together with the CE mark in Figure 10-4, which signifies that restrictions may be applicable in certain member states.



Figure 10-4: The CE mark and the alert symbol

Note that terminal equipment that only transmits under the control of a network can be classified as Class 1 (subclasses 1.9 to 1.17) [297]. This caveat might enable SDR user equipment in client/server networks to operate in non-harmonised frequency bands. This might prove to be important for the implementation of spectrum liberalisation and certain CR applications. This needs to be clarified.

SDR implementing Class 2 equipment and the R&TTE directive are on a potential collision course. The R&TTE directive requires that the relevant regulatory body is notified 28 days in advance of any product that operates in non-harmonised frequency bands being commercially released. Note that if SDR is seen as an enabling technology for spectrum liberalisation, this by definition implies the introduction of SDR equipment into non-harmonised frequency bands. The directive also devolves responsibility for ensuring that a product meets the necessary performance requirements to the manufacturer. Both of these might prove to be contentious issues.

The full potential of SDR will be realised when radio equipment can be reconfigured on the fly with the latest configuration data. It is quite reasonable to assume that software updates will be released after the hardware. Again we return to the question as to whether it is the hardware or the software that requires type approval. In order to comply with the 28-day notification period it *has* to be the software. The hardware will need some form of approval of course, but against what it should be evaluated is not clear. Who has responsibility for SDR type approval under the R&TTE directive? It is the software that will dictate the function of an SDR, but the software vendor will not want to take responsibility for the performance of the hardware, especially if, in the long-term, methods are devised to parameterise the signal path and enable higher-level, non-platform-specific software radio to be defined.

UNCLASSIFIED

The R&TTE directive represents a significant milestone in the regulation of telecommunications equipment in Europe. However, a review to better support the realisation of SDR might be necessary. Failure to do so might deter innovation and the full realisation of the benefits of SDR.

10.9 SDR and Spectrum Trading and Liberalisation

Historically, the process of regulating spectrum usage has been to allocate sections of the RF spectrum for particular uses at the international level. Harmonisation and agreement at the international level is generally required to minimise and manage cross-border interference, foster the adoption of globally recognised telecommunications standards and support the globally recognised radio systems required for maritime, aviation and navigational purposes. Both licensed (e.g., cellular) and unlicensed (e.g., industrial, scientific and medical (ISM)) spectrum bands are defined. Within licensed spectrum bands, licences are assigned to spectrum users at the national level by the relevant spectrum regulator. In the UK, this task falls to Ofcom. This process of the national regulator controlling access to licensed spectrum is commonly referred to as a ‘command and control’ approach to spectrum management. In 2000, over 95% of the usable radio spectrum in the UK was managed by this approach as shown in Figure 10-5 [8].

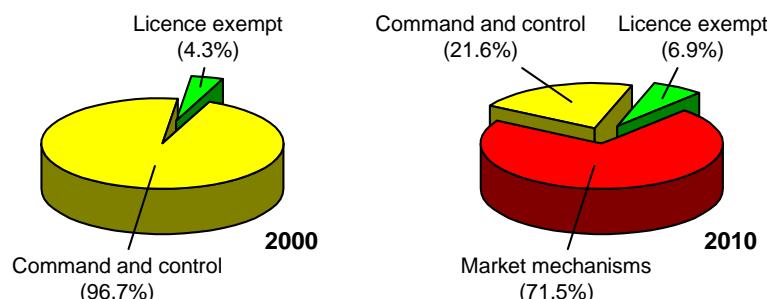


Figure 10-5: The balance of spectrum management in the UK in 2000 (left) and 2010 (right) [8]

Whilst the global harmonisation of spectrum has its advantages, it generally leads to inefficient use of the available radio spectrum; some bands (e.g., cellular bands) are heavily congested whilst other parts of the spectrum are largely ‘empty’. Therefore, given the ever increasing demand for new spectrum from users, the traditional approach of imposing stringent restrictions on the use of a particular RF band and the general availability of addition RF spectrum is rapidly becoming unacceptable.

Following the findings of first Cave [7] and, following on from this, Ofcom’s spectrum framework review [8], Ofcom is keen to relax many of the restrictions placed on the users of licensed spectrum and to promote the concepts of spectrum trading and spectrum liberalisation. Both concepts aim to improve the utilisation of the available RF spectrum.

UNCLASSIFIED

Spectrum trading provides spectrum users with an incentive to optimise their spectrum utilisation. By making the most efficient use of their spectrum they then have the option to gain extra income by leasing (or selling) the rights to use unused spectrum to third parties. Conversely, in areas of high traffic density operators have the potential to lease additional spectrum in order to meet the local traffic density demands. Given the high financial value typically associated with RF spectrum, there will thus be a clear incentive for operators to optimise spectrum utilisation. With spectrum trading, detailed spectrum management is devolved to the spectrum users and 'market mechanisms' will drive the need for commercial users to make efficient use of their allocated spectrum.

Ofcom's aim is for over 70% of the radio spectrum to be managed by market mechanisms by 2010, with management through command and control falling to just over 20% of the usable spectrum, as shown in Figure 10-5 [8]. Note that, not only will the adoption of market mechanisms to manage spectrum usage encourage efficient spectrum use, it will also significantly reduce the effort required on the part of Ofcom to police the spectrum.

Spectrum trading refers to the trading of spectrum rights between users. Spectrum liberalisation removes the restrictions that might otherwise inhibit the innovative use of technology to improve spectral efficiency beyond 'simple' network optimisation.

The benefits of spectrum liberalisation are many. Again, spectrum liberalisation has the ability to help maximise spectrum utilisation. It gives the user the option of adopting new, innovative radio technologies that can be optimised to the quality of service requirements of the user, thereby making better use of the available spectrum. Moreover, it gives users the ability to optimally 'mix-and-match' radio technology characteristics to deliver a range of service types. This 'mix-and-match' approach might constitute the beginnings of a CR system. Depending on restrictions imposed by existing, internationally agreed harmonised frequency bands, spectrum liberalisation also presents an opportunity to have different radio technologies and services sharing common spectrum.

An example scenario that might result from the introduction of spectrum trading and liberalisation is shown in Figure 10-6. One operator provides a cellular telecommunications service. Another operator broadcasts DVB-T transmissions. Each operator has sole transmitting rights to separate RF bands. Additionally, however, an agreement has been negotiated between the two operators to share an additional band of RF spectrum. Under this agreement, the cellular operator has access to the shared spectrum between 6 am and 6 pm to add the extra capacity required to support business users and high bit rate data services on its network. During the evenings, the spectrum is taken over by the DVB T broadcaster to add additional primetime TV channels to its customers. This arrangement would provide both operators with additional capacity when they most need it whilst significantly reducing overall spectrum costs.

UNCLASSIFIED

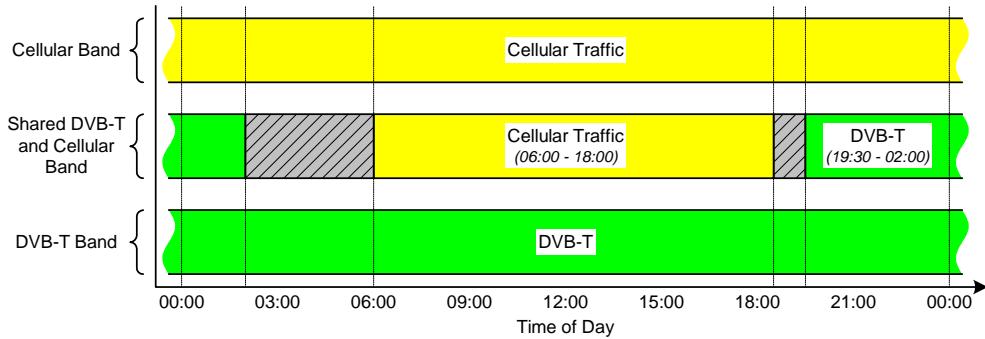


Figure 10-6: An example of spectrum trading and liberalisation between cellular and DVB-T bands

A second spectrum trading and liberalisation example is shown in Figure 10-7. Here, a certain amount of RF spectrum is allocated to the public safety networks. Some of the time, the public safety spectrum might be underutilised. Through careful management of its carriers, the public safety network operator might lease unused spectrum to third parties such as commercial cellular network operators on a real-time basis. Under this arrangement, cellular operators have a relatively low cost path to increasing nominal network capacity. To the public safety network operator, leasing out temporarily unused spectrum adds an additional source of income. It is acknowledged that such an arrangement would not be without technical challenges. It would be essential that public safety spectrum leased to other users could be instantly and reliably released for public safety traffic, e.g., in the case of a major incident. Releasing cellular spectrum for public safety use during major incidents might also be an attractive option. An efficient method of negotiating and billing the leasing of surplus spectrum would also be required. However, this example is indicative of the kind of innovative solution to improving overall spectral efficiency that might become reality following the introduction of spectrum trading and liberalisation.

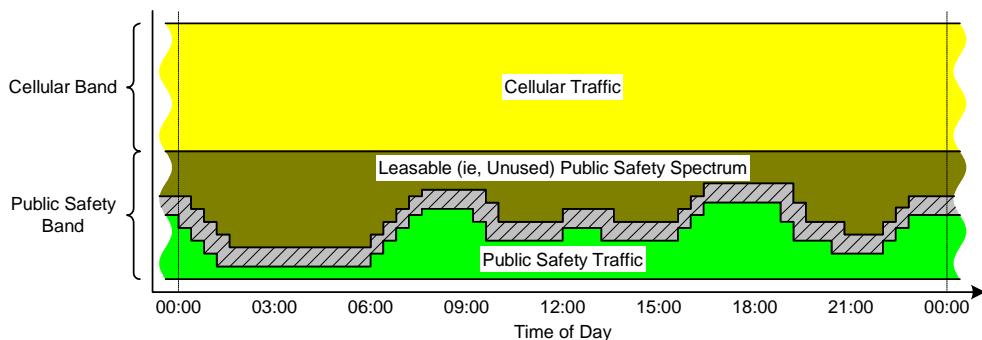


Figure 10-7: A simplified example of spectrum trading and liberalisation between cellular and public safety band users

UNCLASSIFIED

A final example of spectrum trading is that where a spectral band is reused on a geographical basis. Thus, spectrum used by a service provider operating in London, might be reused by a different service provider in Manchester. It will be the responsibility of the relevant parties to negotiate on issues such as inter-operator interference.

So where do SDRs come into this? SDRs will potentially facilitate the rapid switching between different radio technologies and RF bands. Thus, SDR will enable the spectrum users to switch operation from one frequency band and/or technology to another rapidly, without having to acquire new radio equipment or discard existing equipment.

There are no regulatory issues regarding SDR that are specific to the concepts of spectrum trading and/or spectrum liberalisation. It is important to note that SDR, given its potential to offer flexible, reconfigurable radio platforms, may well prove to be the enabling technology that leads to the widespread acceptance of these concepts. However, there is a possible chicken-and-egg scenario that might need to be kick-started when considering spectrum trading and liberalisation and SDR. Whilst SDR might ultimately prove to be the enabling technology for these concepts, it is possible that it will require the general adoption and acceptance of these same concepts to drive the development of SDR.

Note that although spectrum trading and the concept of using market mechanisms to manage spectrum usage are compatible with the current international regulatory framework of harmonised frequency bands, it is doubtful that the concept of spectrum liberalisation is. Many of today's harmonised frequency bands dictate not only the purpose for which a particular RF band should be used but also the specific radio technology (or technologies) that may be deployed within it. It is acknowledged that, even with the advent of spectrum liberalisation, international harmonisation is still required to a certain degree. Indeed, the continued global harmonisation of bands used for navigation, aviation and maritime are essential. However, in many cases, rather than allocating spectrum to a particular radio technology, an ultimate goal might be for spectrum to be allocated internationally to broad service categories, e.g., mobile, fixed wireless links, etc., as recommended by Cave [7]. Such a vision will only become reality through negotiations at the international level.

10.10 Conclusions

In this chapter we have considered regulatory issues in some detail. SDR has the potential to revolutionise radio design. In particular, the realisation of field reprogrammable SDRs will open the door to the development of a wide range of innovative new radio technologies that can be optimised and refined to precisely meet the service requirements of the user. Thus, SDRs might ultimately prove to be a key enabling technology for the concepts of spectrum trading and spectrum liberalisation, which Ofcom is keen to pursue with the view of better managing RF spectrum in the UK and promoting the more efficient utilisation of the available spectrum.

For non-field programmable SDRs (i.e., SDRs that are configured at the time of manufacture and cannot be reprogrammed in the field) there are no significant regulatory issues as this type of equipment differs from more conventional radio equipment only in the manner in which the signal path is implemented.

UNCLASSIFIED

The regulatory procedure currently implemented in the UK, specifically the R&TTE Directive 1999/5/EC, is not particularly accommodating of field programmable SDRs. The current regulations, which make the manufacturer responsible for ensuring type conformance, do not appear to allow for the situation where one party manufactures SDR hardware and another party develops the software. Furthermore, the need for a 28-day notification period before products that operate in non-harmonised bands can be marketed might prove contentious; if the equipment has to be physically marked then this effectively prevents new software configurations from being downloaded to SDR hardware that is already in circulation. A solution to product marking might be to resort to some form of electronic labelling as permitted by the FCC in the US.

A possible loophole that might allow SDR user equipment in client/server networks to operate non-harmonised radio technologies and/or in non-harmonised frequency bands is that equipment that only transmits under the control of a network can be classified as Class 1 equipment. Class 1 equipment can be deployed without restriction. The validity of this approach needs to be clarified.

The FCC has demonstrated a commendable desire to evolve its regulatory processes to embrace the flexibility of SDRs and CR. In particular, by defining a Class III permissive change for SDR software changes, and by allowing the use of electronic displays to display the FCC authentication number, the FCC has streamlined the type approval process and paved the way forward for the use of field programmable SDR in the US.

Long-term, however, there is still a danger that the regulatory procedures will inhibit utilisation of the full potential of SDR. For example, the FCC require conformance testing of all hardware/software combinations before software updates can be released. Whilst this is certainly the safest route, it might discourage the development of a high-level, innovative method of defining new SDR configurations. This, coupled with the restriction that only the original application can apply for Class III permissive changes, may deter third parties developing innovative SDR applications that may be targeted at a wide range of SDR hardware.

This said, the FCC acknowledges that SDR and CR as realisable, practical concepts are still in their infancy and fully expects further policy reviews as the technologies mature and become better established.

11 Commercial Drivers



By Tim James, Multiple Access Communications Ltd.

11.1 Introduction

In this chapter we examine the key factors that have driven the introduction of software defined radio in the commercial arena in recent years. We also speculate on the commercial drivers that are likely to define the progress that will be made in this area in the years to come.

When examining these drivers, it becomes clear that different types of organisation (e.g., manufacturer, network operator) have quite different reasons for adopting SDR technologies. In the following sections we consider the commercial drivers from a range of different perspectives. Many of the views expressed in this section have been derived from a workshop held on 26 January 2005 with some of the key 'stakeholders' from the SDR community, including network operators, equipment manufacturers and semiconductor houses. We would like to take this opportunity to thank these organisations for their contribution to this project.

11.2 A Manufacturer's Perspective

There is a range of different manufacturers involved in the design, development and production of radio equipment. As we start to consider the motivation for each type of manufacturer to adopt SDR technologies, we discover once again that different types of manufacturer will have different reasons for adopting SDR. In the following sections we consider the factors driving the adoption of SDR in each type of manufacturer.

11.2.1 Infrastructure Vendors

Cost has always been a key driver in the development and production of telecommunications infrastructure equipment, but in recent years its importance has increased significantly such that it is perhaps the single, dominant factor on which infrastructure vendors compete for business. To understand why, let us consider the cellular industry as an example. Mobile number portability (MNP) removed one of the last significant barriers presented to a cellular subscriber wanting to move between networks and this means network operators have to fight even harder to increase or even retain market share. However, in a mature market, how can network operators differentiate their product offering from their competitors? In the early days of network deployment, operators could compete based on network coverage and quality, but now in most cases all networks offer similar levels of coverage and call quality and this battleground has largely been eroded. These days the key battlegrounds are handset range and availability and the cost of the service to the end user. The latter of these two factors means that cost has become a dominant factor within the industry.

UNCLASSIFIED

The 3G spectrum auctions left many network operators with little financial capital to purchase and deploy their 3G networks and this led them to apply significant price pressure to the infrastructure vendors. In an attempt to decrease the cost of their equipment, the infrastructure vendors have turned towards SDR, which will impact on the equipment costs in the following ways:

- a) *Decreased equipment development costs.* SDR can decrease the development costs associated with a particular piece of equipment in a number of ways. Firstly, by using programmable devices, e.g., FPGAs and DSPs, the costs associated with the development of an ASIC can be eliminated.

Secondly, development teams become familiar with using particular programmable devices and the associated suites of development tools. This means that the costs associated with getting ‘up-to-speed’ when development starts on equipment based on a new technology standard (e.g., UMTS, WiMax) can be minimised.

As an example, picoChip Designs Limited have produced reference designs for the 3GPP UTRA technology standard, including HSDPA, and the WiMax (IEEE 802.16) technology standard. Both of these reference designs are based around hardware platforms that utilise two of picoChip’s PC102 programmable signal processing devices.

Therefore, a development team that is already familiar with these devices from, say, a UMTS equipment design project, will have a significantly reduced learning curve when it comes to designing WiMax-based equipment.

If we compare this to the scenario where custom and semi-custom devices are used that are targeted at a specific technology standard, when development starts on equipment based on a new technology standard, the development teams must become familiar with the new technology-specific devices.

Finally, as a development team generates an increasing number of equipment designs, it will not only accumulate valuable experience and knowledge in the use of a particular set of devices, it will also accumulate sets of building blocks (e.g., filters, demodulators, channel coders) that can be reused in new designs. This idea of intellectual property blocks is an important factor in the development of SDR and we might expect companies to spring up who’s primary business is the development and sale of IP blocks for programmable devices.

UNCLASSIFIED

- b) *Decreased hardware platform costs.* In scenarios where a particular hardware platform can be used across a range of different products, the costs associated with hardware platform design and development can be minimised. The use of programmable devices means that hardware platforms become more flexible and the potential to reuse hardware platforms across a range of products and radio technologies is increased. Once again, if we contrast this with the use of custom and semi-custom devices that are targeted at particular technologies, the hardware platform must be redesigned to support each new technology. Hardware development costs can be decreased still further if the hardware platform can take the form of a more general purpose off-the-shelf product. Vanu Inc. have recently developed a base station conforming to the GSM standard based on an off-the-shelf HP ProLiant server with two 2.8GHz Intel processors [298]. The server is also used to implement the base station controller (BSC) functionality and the radio front-end is based on the off-the-shelf ADC Digivance radio transceiver. In this example, the infrastructure manufacturer is not required to perform any hardware development and all development work is performed in the software domain. Also, since the hardware that is used has application in a wide range of different areas, the costs will be significantly lower than those associated with equipment that is specifically designed for the radio market.

In addition to these cost advantages, programmable devices also provide a range of other benefits to equipment manufacturers that will drive the adoption of SDR. These benefits include:

- c) *Faster time-to-market.* Equipment vendors can gain significant advantage if they can bring their products to market more quickly than their competitors. In addition to decreasing the development costs, the use of programmable devices can also decrease the development timescales and, hence, the time-to-market can be significantly reduced.

UNCLASSIFIED

- d) *Ability to track changing wireless standards.* Since the development of the GSM system, radio standards have typically been released in phases. The initial (Phase 1) release of the standards is generally used to provide a stable description of a system with a minimum set of features. Once this initial release has been frozen, equipment manufacturers can produce equipment in the knowledge that it conforms to a stable, recognised standard. The second release of the system standards (Phase 2) is used to introduce new features and also provide fixes for problems identified in the Phase 1 release. In the case of the GSM system, there were only two full releases of the system standards. However, the system continues to evolve and the appropriate parts of the system standards are released as and when required to accommodate new features. This process has been termed 'Phase 2+'. In the case of the UMTS standard, the initial release of the system standard was termed 'Release 99' and this was finalised in March 2000. The subsequent release, Release 4, was finalised in March 2001 and Release 5 was finalised in June 2002. The next release, Release 6, is due to be finalised in March 2005. From this discussion it is clear that standards evolve considerably throughout the lifetime of a system and the equipment manufacturers must adapt their designs to track this evolution since they cannot afford to wait until the standards have been frozen to commence work on their equipment designs. The use of programmable devices provides equipment manufacturers with the flexibility to track the changes in the relevant system standards as they move towards finalisation and rapidly bring equipment to market once the standards are frozen.

- e) *Post-manufacturer bug fixes.* Although all equipment will be thoroughly tested before it leaves the equipment manufacturer, with complex communications system, there is always the potential for problems (i.e., equipment 'bugs') to become apparent once the equipment has been installed. The use of programmable devices increases the potential to fix bugs in equipment once it is in the field by means of a software download, rather the physical replacement of parts of the equipment. In general, the software download approach is likely to offer a far more cost effective solution when compared with replacing equipment, even if an engineer has to visit the equipment site to perform the download. However, in situations where software downloads can be performed from a remote location, the costs associated with fixing a problem could be orders of magnitude smaller than those associated with visiting all of the equipment sites and replacing the faulty equipment. This is particularly the case in cellular networks where it may be very difficult to gain access to microcell and picocell sites.

UNCLASSIFIED

11.2.2 Terminal Manufacturers

In large scale mobile radio networks, the business of designing, manufacturing and selling terminal devices (e.g., mobile phones, laptop wireless cards, wireless personal digital assistants) is very different to the business of designing, manufacturing and selling infrastructure equipment. Firstly, the end customers are very different and, hence, they have very different requirements and priorities. The end customers for terminals are mainly consumers and the key factors that drive the success of terminal sales, particularly mobile phones, are cost, size, weight, battery life, cosmetic design and the number of features it contains (e.g., built-in camera, MP3 player, radio). Secondly, the number of terminals produced and sold will be orders of magnitude greater than the number of infrastructure products sold (e.g., base stations). For example, there are currently around 35,000 cellular base stations in the United Kingdom [299] serving over 56 million cellular subscribers [300] each with at least one mobile phone. Over 650 million mobile phones were sold worldwide in 2004 and 200 million of these were made by Nokia, the largest mobile phone manufacturer in the world [301]. Let us now examine how these two key aspects of the terminal business may affect the uptake of SDR techniques within these devices:

- a) *End user requirements.* Many of the factors driving the design of mobile terminals tend to oppose the use of programmable devices. For example, programmable devices tend to be more power ‘hungry’ than dedicated devices and this will decrease the battery life of the terminal and also increase the potential weight and size of the product. Therefore, terminal manufacturers are likely to steer clear of flexible, programmable device unless there are some overwhelming factors that support the use of these devices. If we examine the range of terminal chip sets currently on the market, then they are generally targeted towards one or two specific technologies, rather than being sufficiently flexible to support a range of radio technologies. It appears that the main driver behind chipset development is feature integration, i.e., providing terminal manufacturers with the ability to support as many features as possible with as few chips as possible.

- b) *Volume sales.* The fact that terminals are generally produced in very large numbers means that terminal manufacturers can justify the up front investment in custom devices such as ASICs, which offer lower power, smaller footprint solutions when compared with programmable devices. This will tend to restrict to use of programmable devices in high volume products. In lower volume products (e.g., GSM-R, TETRA, dual-mode cellular/public safety terminals), the use of programmable devices may be more prevalent, since the product volume may not justify the costs associated with developing a tailored ASIC.

UNCLASSIFIED

In addition to the points discussed above, a number of other interesting views were expressed during the Stakeholders' meeting. It was felt that as new technologies emerge, the available programmable devices are not necessarily capable of supporting the new technology within a terminal design. Therefore, there is little alternative in the initial terminal designs to the use of ASICs. However, as the capabilities of the programmable devices improve, then they start to offer an alternative in the design of later generation terminals. The recent announcement by Nokia that it is going to use Texas Instruments' single chip solution for future handsets [302] was also seen as a significant move amongst the Stakeholders. Nokia has traditionally developed its own chipsets for its terminal designs, but it is finding this approach to be too costly for high volume, entry-level terminal markets (e.g., China and India). By using the Texas Instruments solution it can decrease its development costs and provide lower cost entry level products. Although this announcement does not represent a move towards the use of SDR, it does signify an acceptance by terminal manufacturers of third-party chipsets and this could pave the way for greater use of programmable devices in the future. It was also felt among the Stakeholders that terminal manufacturers are becoming less concerned with differentiating themselves through the core terminal technology and more emphasis is being placed on product differentiation at the man-machine interface (MMI) level. In the longer term, this may encourage the use of a common set of flexible, programmable radio platforms that can be applied to a range of different technologies with each terminal manufacturer customising their own product at the MMI level.

During the Stakeholders' meeting some time was also spent discussing the issue of using SDR techniques to 'future-proof' terminals against changes in radio standards. The view was expressed that terminal manufacturers would not necessarily be in favour of producing terminals that could be upgraded to extend their useful life, because this would decrease the amount of terminal 'churn' within the subscriber base and, hence, decrease the level of terminal sales. However, after discussion amongst the Stakeholders, it was felt that this was only a minor consideration for the terminal manufacturers since terminal churn was much more likely to be influenced by the cosmetic design of the terminals and the latest terminal 'fashions', than being influenced by the upgradeability of the terminals.

11.2.3 Semiconductor Houses

Digital semiconductor devices used within radio equipment appear to fall into three main categories, namely, general programmable devices, wireless-specific programmable devices and technology-specific devices.

The generic programmable category includes devices such as DSPs and FPGAs. These devices can be used in a wide range of applications and they are not specifically designed for use in radio transmitters and receivers. Devices of this nature are currently being used extensively in commercial radio equipment where power consumption is not a significant limitation, e.g., in fixed infrastructure equipment such as cellular base stations.

The wireless-specific programmable category includes programmable devices that are specifically designed for SDR applications. However, these devices are not targeted at particular radio technologies and they are made sufficiently flexible to support a range of different signal processing operations.

UNCLASSIFIED

Finally, the technology-specific category includes devices that have been designed with a particular radio interface technology in mind and, although they may include a degree of programmability, they cannot generally be adapted to support other technologies. An example of these devices is the TCS3500 chipset developed by Texas Instruments, which has been specifically designed for use within terminals supporting the EDGE technology [303].

The fundamental business of a semiconductor manufacturer is to sell its chips into as many products as possible and, when we consider the commercial drivers behind the adoption of SDR within the semiconductor industry, we must ask ourselves how SDR will help a manufacturer to sell more devices. In many cases the semiconductor manufacturers will be attempting to track the requirements of the organisations further along the supply chain (i.e., manufacturers, operators, end users). Therefore, if these organisations are demanding products that are more suitable for SDR designs, then the semiconductor manufacturers will adapt their products to suit this demand.

Some companies have identified an opportunity to develop new products that compete with the established players in this area. One example of such a company is picoChip Designs Limited (picoChip), who has developed a parallel processing device that is specifically targeted at wireless infrastructure products. The company was founded in 2000 with the aim of providing solutions for next generation wireless systems. Its picoArray™ parallel processing device has the potential to replace ASIC, FPGA, DSP and embedded processor devices and picoChip offers reference designs for the HSDPA technology and the WiMAX technology based around its products. Discussions at the Stakeholder meeting indicated that semiconductor customers increasingly expect to buy complete systems, including software solutions, rather than just simply buying chips and developing their own software designs. In fact, in the case of picoChip, most of the development effort within the company is focussed on developing software reference designs based on its semiconductor products, rather than on development of the semiconductor devices themselves.

More traditional semiconductor companies are attempting to gain a larger share of the wireless market by showing the manner in which general purpose DSPs and FPGAs can support the new range of wireless technologies that are currently being defined. For example, Xilinx have generated reference designs for various parts of the UTRA FDD standard based on its FPGAs, including the random access channel detection process [304], the uplink channel searcher function [305] and a HSDPA base station receiver [306]. DSP manufacturers are also attempting to gain a larger share of the wireless market by developing faster, more powerful devices that can compete with ASICs and FPGAs, e.g., the TigerSHARC® processor from Analog Devices [307].

To summarise, it appears that SDR presents opportunities within the semiconductor industry for both new and established companies. If SDR becomes a strong requirement for organisations further along the supply chain, then semiconductor manufacturers will support these requirements as much as possible to gain market share. Also, SDR may allow semiconductor manufacturers to consolidate their product portfolio around a limited set of flexible, programmable devices, rather than a wide array of devices targeted at different technologies. This will allow product development costs to be spread across a wider range of end applications, thereby potentially decreasing the cost of the products to the end users.

UNCLASSIFIED

11.2.4 Test Equipment Manufacturers

In addition to the terminal, infrastructure and semiconductor manufacturers, it is also important to consider another key manufacturer of radio equipment, the test equipment manufacturer. The test equipment industry has been exploiting SDR techniques for many years. With modern wireless test equipment, it is common for customers to purchase a ‘main frame’ test equipment platform with a number of specific optional features for testing a particular technology, as opposed to purchasing a test set that is designed solely for testing one radio technology. The Rohde & Schwarz CMU200 Universal Radio Communications tester is a good example of a flexible test equipment platform [308]. This product can support a wide range of technologies, including GPRS, EDGE, GSM, Bluetooth wireless technology, TDMA (IS-136), AMPS, W CDMA and CDMA2000 based on different software and hardware options. The SDR approach has a number of benefits from the perspective of a test equipment manufacturer, including the following:

- a) *Lower development costs.* In situations where the testing of new radio technology standards can be based on an existing hardware platform, the development time and costs associated with producing test equipment based on these new standards can be significantly reduced compared with a situation where a new hardware platform must be developed. In addition to decreasing the overall cost of the equipment, this allows the test equipment manufacturer to bring their product to market more rapidly. Also, the hardware development and production costs can be decreased in situations where a manufacturer can consolidate its product portfolio around a limited set of hardware platforms.
- b) *Increased customer loyalty.* Another key benefit of the ‘main frame’ approach is improved customer loyalty and decreased customer churn. Where a customer has invested in a flexible test equipment platform they are more likely to return to the manufacturer of the platform to purchase an upgrade for a new wireless technology than to investment in a completely new hardware platform.

11.2.5 Software Houses

SDR presents an opportunity for companies to develop and sell software to run on SDR platforms. Based on our investigations and the discussions at the Stakeholders’ meeting, there appears to be two main types of software product opportunities within the SDR market. Firstly we have IP blocks, which are software modules or device configuration files that can be used to implement parts of a radio system on a programmable device. Most device manufacturers provide their own IP blocks and many also have set up developer networks of third party companies who supply IP blocks based on the manufacturers’ products. Examples of these networks include the Texas Instruments Third Party Network [309] and the Xilinx AllianceCORE partner programme [310]. As programmable devices find their way into an increasing number of radio products, then we can expect the market for IP blocks to increase and more companies to enter this arena. Therefore, we can say that there are strong drivers towards the adoption of SDR techniques for suppliers of IP blocks.

UNCLASSIFIED

The second type of software providers is those organisations who develop complete software applications to run on an SDR platform. In contrast to the IP block provider, these organisations could sell their products to the end user of the SDR device rather than the SDR manufacturers. At present the market for this type of software is practically non-existent, because SDR platform manufacturers do not generally allow third party software developers access to their products, other than perhaps at the design stage. However, as the SDR market develops, we may see more open SDR platforms and this is likely to stimulate growth in the number of companies developing third party applications. The growth in this area will be influenced by a number of factors, including the willingness of the SDR platform manufacturers to open up their platforms to third parties and the regulations that control the use of third party software on SDR platforms.

11.3 A Network Operator's Perspective

Having considered the commercial factors that will drive equipment manufacturers towards or away from the use of SDR techniques, we now consider the commercial drivers for the use of SDR within network operators, i.e., organisations who deploy and operate large-scale public access or private radio systems (e.g., cellular network operators, public safety network operators, fixed wireless access network operators). When we consider these drivers it becomes clear that they can be divided into two broad categories based on the type of SDR being considered. One category of drivers relies on the use of programmable devices in the equipment and these do not require the equipment to be reconfigurable after manufacture. The second category relies on the use of programmable devices in the equipment and the ability to reconfigure the equipment once it has entered operation, i.e., post manufacture. We will consider each of these categories individually below.

11.3.1 Commercial Drivers for the use of Programmable Devices in Equipment

As we have already discussed, the use of programmable devices is likely to lead to decreased equipment costs, particular in relatively low volume infrastructure equipment. Therefore, given the significant cost pressures facing network operators in today's telecommunications environment, this is a significant driver towards the use of programmable devices in the design of infrastructure equipment. A second driver towards the use of programmable devices within infrastructure equipment is the ability to support a number of different radio interfaces within the same equipment. AirNet Communications Corporation has developed a product that supports both a cellular radio interface (e.g., GPRS or W-CDMA) and a broadband point-to-point radio interface [311]. The product is marketed as a 'backhaul free' base station in that the same equipment can be used to provide cellular radio coverage and also an interconnection to the rest of the radio network via the point-to-point radio link, without the use of a separate point-to-point microwave link. In this example, part of the 'mobile' spectrum is used to support the point-to-point links, thereby removing the need for the network operator to acquire additional spectrum to accommodate the backhaul links.

UNCLASSIFIED

This technique could also be used to support ‘multi-protocol networks’, i.e., networks that consist of multiple air interface standards. Equipment of this nature is particularly relevant at the present time as network operators around the world roll out 3G networks alongside their existing 2G networks. Using the software radio approach it would be possible to support the different air interface standards within the same hardware platform.

11.3.2 Commercial Drivers for Post-Manufacture Reconfigurability

Having considered the commercial drivers associated with the use of programmable devices in the manufacture of infrastructure equipment, we now move on to examining the drivers for the use of reconfigurable equipment within large-scale mobile telecommunications networks. One of the main attractions of reconfigurability is that it provides a means to ‘future-proof’ the investment made by network operators in infrastructure equipment. As the radio standards evolve, the installed equipment can be modified to accommodate any changes by means of remote software downloads, rather than more costly hardware upgrades, which require physical visits to each base station location. The ability to upgrade equipment remotely is becoming increasingly important with the proliferation of base station equipment since the costs associated with sending an engineer to every site to perform a hardware upgrade increase as the number of base stations increase. In some cases it is also difficult to gain access to the base stations (e.g., in-building picocellular base stations) and, again, this makes remote software upgrades an attractive option.

However, it is important to note that SDR technology has not evolved sufficiently to ensure that the installed SDR hardware platform can support any future radio interface by means of a simple software upgrade. Therefore, although SDR base stations may be able to accommodate small changes to the radio interface standards, they are unlikely to be able to cope with major changes or a complete switch of technology at the present time, e.g., from CDMA to OFDMA. It is also important to note that the network operators must weigh up the costs and benefits of installing a flexible SDR hardware platform and trade-off the additional costs associated with more flexible equipment against the potential costs savings for future equipment upgrades.

Another driver for the use of reconfigurable equipment in large-scale radio networks is the ability to rapidly deploy new technologies or deliver new services. This driver is associated with the previous one in that it is much easier to upgrade equipment by means of a remote software download than it is to send an engineer to each equipment site to perform a hardware upgrade. It is also potentially easier to fix ‘bugs’ in the equipment once it has been deployed if the equipment can be reconfigured once it is in the field.

UNCLASSIFIED

Another important driver towards the use of reconfigurable radio platforms is the ability to gradually migrate between different radio interface standards. If we look back to the migration between the first generation (1G) TACS networks and the 2G GSM networks in the UK, there was a gradual re-allocation of spectrum from TACS to GSM. The equipment associated with these two technologies was quite different and the migration process consisted of physically removing carriers associated with the TACS technology and replacing these with carriers associated with the GSM technology. As we move from 2G to 3G technologies, the issue of spectrum ‘re-farming’ has again arisen and a similar process may occur as GSM spectrum is given over to 3G technologies. Using SDR, the base station platform could be sufficiently flexible to support both GSM and 3G technologies and the migration between different radio interface standards could be performed by means of a remote software change within the base station. This would allow the network operators to tailor the split between GSM and 3G spectrum in each area to the traffic demands made on each technology.

11.3.3 Commercial Drivers for Reconfigurable Terminals

In addition to the drivers associated with the use of programmable devices in infrastructure equipment, we also considered the factors that might drive a network operator to encourage the adoption of reconfigurable terminals. One of the key drivers is the ability to deploy new technologies or services within an existing terminal base. An example put forward by Eduardo Ballesteros and Carlos Martinez of Telefónica [312] involves the deployment of a new voice codec as a means of capacity management. In a situation where a cell site is overloaded, the network operator could choose to deploy a new lower rate speech codec for use within the cell, thereby decreasing the cell loading and allowing more users to be supported. This is similar to the approach used in the GSM system with the development of the half-rate codec. However, with reconfigurable SDR terminals, the new speech codec can be deployed very rapidly and it can be downloaded only to terminals in areas of congestion.

Another potential driver within network operators for reconfigurable terminals is the ability to interoperate with a wider range of devices. For example, if a customer of a GSM network in Europe travels to the United States, then with a tri-band or quad-band GSM terminal, they will be able to roam onto the GSM networks in the United States (subject to suitable roaming agreements). However, in addition to the GSM networks, there are also extensive cdmaOne and iDEN networks that could provide the traveller with better coverage. If the user had the ability to reconfigure his/her handset to support a chosen radio interface, then he/she would potentially have a much wider choice of roaming options than with a GSM-only terminal. This approach could be used by operators to encourage roaming users onto their networks. In addition, it could be used to move subscribers across to new radio standards as they become available.

11.4 A Consumer's Perspective

Finally in our examination of the commercial drivers that might encourage the adoption of SDR technologies, we consider the perspective of the consumer. Once again, the factors that affect his/her decision to purchase a particular terminal are quite different to the factors that affect the network operators' decision to purchase a particular piece of equipment. Based on our discussions with the various Stakeholders, it appears that the key factors in the consumers' purchasing decisions are the cosmetic appearance of the product (i.e., is it the latest style?), the size, weight and battery life and the number of features supported (e.g., MP3 player, FM radio, Bluetooth wireless technology). Clearly the cosmetic appearance has little to do with whether or not the product makes use of SDR techniques. As far as the size, weight and battery life of the product are concerned, then this will tend to drive users away from devices based on SDR hardware, since these will tend to lead to bulkier, more power hungry devices. Therefore, it seems that the key driver towards the adoption of SDR from the perspective of the consumer is the ability of the product to support a greater number of features. If devices based on an SDR platform can support more features than non-SDR devices, then this may encourage their uptake within the cellular subscriber base. However, given the price sensitive nature of the consumer market, the costs associated with adding this SDR flexibility to the product must be kept to a minimum and this may prove to be a barrier to its adoption.

11.5 Conclusions

In this chapter we have considered the commercial drivers that could encourage the adoption of SDR techniques within the radio communications industry. We have examined these drivers from the perspective of each of the key stakeholders within the industry and considered the arguments for and against adopting SDR techniques within each part of the value chain. Our main conclusions are that the drive towards the adoption of SDR technologies appears to be strongest within the network operators, where equipment 'future-proofing' and the ability to rapidly introduce new technologies and services are attractive features for SDR. We also see a strong drive towards SDR in the infrastructure equipment and test equipment manufacturers. It appears that the organisations with the weakest drive towards SDR are the terminal manufacturers, where the factors that drive the design of the latest terminals (e.g., size, weight, power consumption) tend to oppose the use of SDR technologies.

Therefore, if SDR is to be a means of supporting spectrum trading and spectrum liberalisation through reconfigurable radio and cognitive radio, it appears that the area of reconfigurable terminals must be addressed and factors must be introduced to encourage the use of SDR in these devices.

12 Assessment of SDR's First Applications and Areas of Deployment



By Julie Fitzpatrick, QinetiQ.

12.1 Introduction

An assessment of software defined radio (SDR) with an emphasis on its application and areas of deployment is considered in this chapter. Attention is focused on the commercial applications of SDR with a review of the various market sectors that SDR may be applicable to, a discussion on SDR products that have begun to emerge and a view on future deployment of SDR in these areas. Epoch estimation for the future applications of SDR is also provided. Included within the discussion are the following:

- Investigation of where SDR has been deployed
- Discussion of the areas where SDR is planned to be deployed in the short term based on the enablers and benefits of SDR identified so far within this study
- Comparison of deployment of SDR across market sectors
- Development of a time line for the deployment of SDR.

12.2 SDR in Cellular Networks

The most obvious market for commercial communications equipment is the cellular network industry and so this is a good place to start when looking at markets where SDR could potentially have an impact. In an industry where standards are continually changing, SDR could provide the opportunity for manufacturers to future-proof their equipment. It could also help them to incorporate additional functionality into their products by supporting a variety of standards.

As SDR devices could be based around a standard programmable platform there is the opportunity to reduce costs. Finally, SDR could offer network operators increased spectral efficiency as the flexibility of SDR opens the way to use more efficient techniques for allocating bandwidth amongst a number of users in a system.

SDR has already been slowly making its way into cellular network products. However, it has not always been promoted as SDR and therefore, in some cases, gone unnoticed. The cellular network products that have been branded as SDR have been slow to be taken up in commercial cellular networks.

This section reviews the introduction of SDR to the cellular industry. It begins by reviewing the history of mobile standards evolution in Europe and the US to set the scene for why SDR might be useful to this industry. It then reviews what SDR cellular products have already been released and looks at the success of these. Next an industry map of the cellular base station industry is shown to identify attitudes towards SDR throughout the supply chain and to draw conclusions on why SDR products have had limited success. Finally, the introduction of SDR into mobile phones is discussed and conclusions are summarised.

UNCLASSIFIED

12.2.1 Standards Evolution in Cellular Networks

The cellular network industry has seen a continuous evolution of mobile standards that operators have been forced to keep pace with in order to remain competitive. This evolution of standards, both in Europe and the US is shown in Figure 12-1.

As demonstrated by the delayed roll out of 3G, replacing network infrastructure to support a new mobile standard is a costly operation and one that shouldn't be entered into until the market realises the benefits of the new service and is ready to embrace it. Getting the funds in place at the right time and developing a roll-out plan that will support the gradual migration of existing users from the old standard to the new is crucial.

SDR is being marketed as a technology that could ease this transition by replacing costly local hardware upgrades with remote software updates. The flexibility of SDR could potentially enable a cellular network of base stations to be upgraded to a new standard via a software download. The base station could also be remotely reconfigured to balance the load between the old and new standards so that the service availability tracked the migration of users to the new standard.

It is important to realise that the ability of SDR to future-proof devices will be limited by the RF hardware of the device and also the processing power of the programmable platform it is run on. While introducing modifications to the same standard, (for example the evolution from GSM to EDGE) might be possible, ensuring that the hardware can support a more complicated standard like 3G, is much more difficult.

UNCLASSIFIED

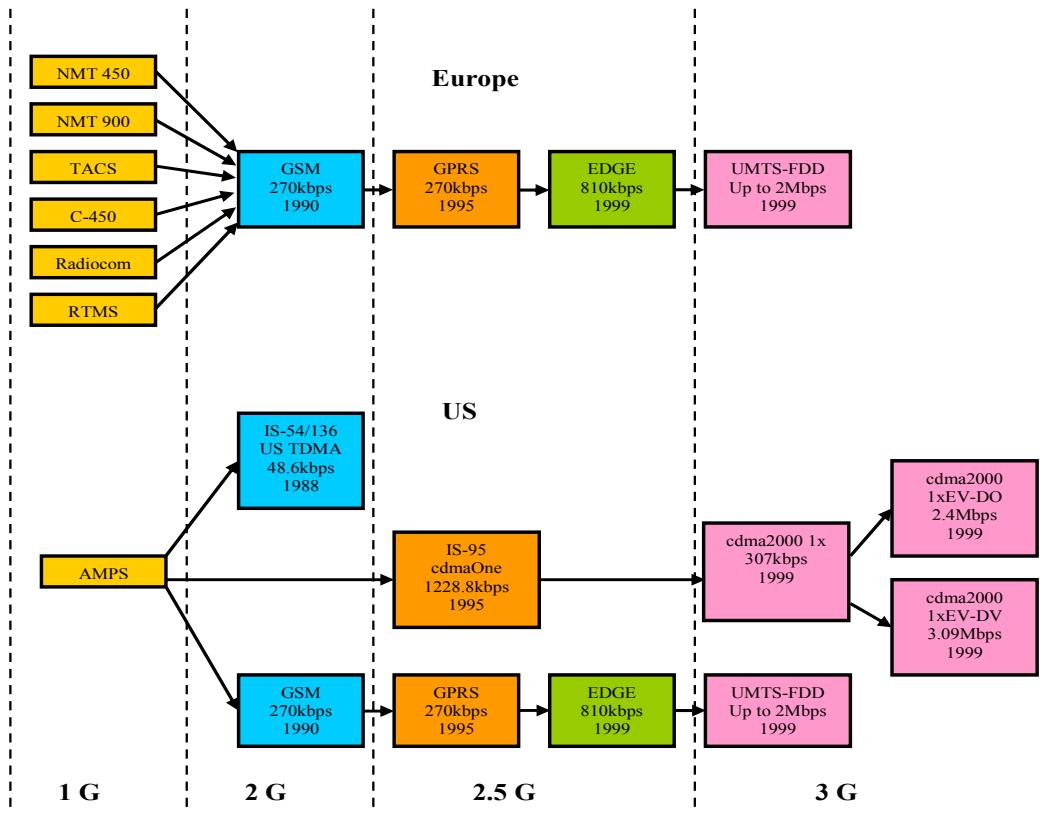


Figure 12-1: Evolution of Cellular Standards in Europe and US [313][314]

In a world of evolving standards, SDR also offers the benefit of faster development times and time to market as it avoids the lengthy process of developing ASICs. Also hardware costs could be reduced as an SDR device is based around a programmable platform that could be used across a variety of products.

An additional benefit of SDR is the ability to support multiple standards. Looking to the future, an SDR mobile phone could reconfigure depending on the environment it was working in. For example, an SDR mobile phone might be configured to GSM initially and then switch to a 3G service when it enters an area with 3G coverage. All this would be transparent to the user with only an improvement in service being observed.

The ability to support multiple standards also gives SDR an added benefit to cellular network operators, particularly in the US. In Europe, it was decided to have a common standards body called the European Telecommunications Standard Institute (ETSI), out of which GSM was born. Worldwide, GSM is by far the dominant cellular standard, accounting for 75% of the world's digital mobile users by the end of 2004 [315].

UNCLASSIFIED

However, the US has adopted a competition driven approach to cellular standards. The result is that the cellular market in the US contains three competing 2G technologies, US-TDMA, cdmaOne and GSM. By January 1999, 120 countries had chosen GSM compared with 24 for US-TDMA and 12 for cdmaOne [313]. Within the US, competition between these three standards is much closer, leading to many US domestic carriers having difficulty deciding which standard to support. Currently, most support US TDMA but, as shown in Figure 12-1, this doesn't have a natural upgrade path to 2.5G and 3G. Some carriers are faced with the option of changing to GSM, cdmaOne or one of the 3G standards [316]. With an SDR solution carriers would not have to choose between standards and would be able to upgrade and balance traffic between services as required.

An additional factor that supports the use of SDR is the entry and increased popularity of Wireless Local Area Network (WLAN) and Wireless Metropolitan Area Network (WMAN) standards. Traditionally, high speed data services like accessing the Internet have been used by fixed users and the mobile market has been focused around voice services. Recently, fixed users have been becoming more mobile by being able to connect their laptops wirelessly into a network via wireless local area network access points in so called hotspots. On the other hand, mobile users have been enjoying increased data rates through 3G.

There is an overlap between the two industries starting to emerge which could lead to increased competition. The ability of SDR to support multiple standards would enable cellular service providers to incorporate and make use of "rival" standards like 802.11 (WLAN) and 802.16 (WMAN) rather than having to compete against them. This relationship is discussed further in Section 12.3.1 which examines the suitability of SDR products for commercial wireless networks outside traditional cellular mobile phone networks.

As pointed out by the SDR Forum's Marketing group [317], it is important that the consumer isn't flooded with numerous confusing mobile standards to choose from. SDR presents the opportunity to offer mobile phone users increased service by utilising a variety of cellular and wireless standards. Crucially, SDR ensures that changes between standards are transparent to the user. This avoids confusion on the consumers' part and also ensures that no one standard must loose out in a cellular standards battle.

12.2.2 Current SDR Cellular Network Products

SDR is a concept that has been discussed since the 1990s but has been slow to be implemented and accepted in commercially available products [318][319]. As discussed in Chapter 11, SDR is most likely to be implemented in base stations first, as size and power constraints make it difficult to use in mobile handsets with current technology. For this reason, commercially available cellular SDR products have been concentrated around base stations.

AirNet has been providing SDR base stations since 1997 but recently has been joined in the market by Airspan (Airspan is a registered trademark of Airspan Networks, Inc.) and Vanu. This section reviews currently available SDR products offered by each of these groups and looks how successfully they have been able to penetrate the commercial cellular market.

12.2.2.1 Typical GSM Architecture

In order to understand where the commercially available SDR products fit into a typical cellular system the architecture of a GSM system is briefly reviewed. Figure 12-2 shows the main components of a GSM network. The main area that SDR has been introduced into is the Base Transceiver Station (BTS) within the Base Station Sub-System (BSS). The Base Station sub-system effectively interprets and relays information from the mobile stations into the Network Sub-System which provides the backhaul link for the network into the Public Switched Telephone Network (PSTN) or gateways to other mobile users on this or other networks. There is also an Operation Sub System that is necessary for monitoring and controlling the network load and capturing billing information.

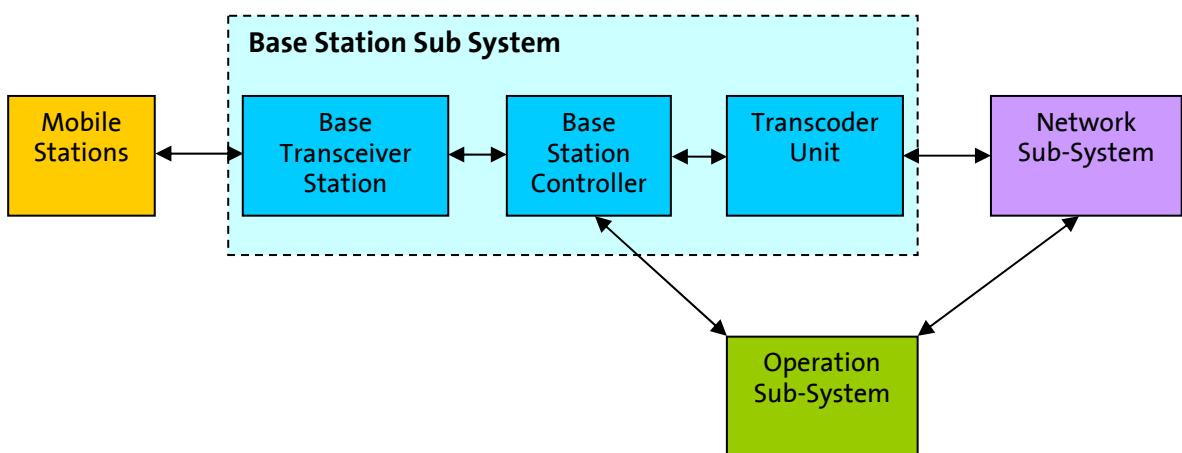


Figure 12-2: Typical GSM Architecture [320]

The waveform implementation occurs in the BTS and so this is where the flexibility of SDR can be best exploited. However, one of the key challenges for SDR BTS providers is that network operators will want a product that is compatible with their existing network components and supports their existing interfaces. Depending on the equipment in the network, these interfaces may be proprietary to one of the major cellular network equipment providers and difficult to obtain details of. This is a considerable barrier to start-up SDR companies.

12.2.2.2 AirNet AdaptaCell® Base Transceiver Station (BTS)

AirNet are a US based company and have been providing SDR base station solutions since 1997 [321]. Their flag ship product is the AdaptaCell® (AdaptaCell is a registered trademark of AirNet Communications Corporation) BTS which is based around a broadband software defined architecture that supports GSM. Software upgrades are available to upgrade the BTS to support GPRS, EDGE and recently 802.16(d). Crucially, this product makes use of the ArrayComm's adaptive antenna array technology which focuses the transmitted power in the direction of the handset. This improves signal quality and network capacity. Such antenna flexibility could also be useful for supporting flexible network architectures which help to overcome the question of how to plan a network of flexible SDR base stations that can remotely switch standards.

UNCLASSIFIED

In terms of interoperability between the AdaptaCell® and existing cellular network infrastructures, the AirNet product range also includes the AirNet Base Station Controller (BSC) and Transcoder Rate Adaptor Unit (TRAU) which comply with the GSM open architecture interface specifications. They have also demonstrated interoperability with some mobile switching centre manufacturers and provide standard T1/E1 links. In addition, they accommodate smaller networks starting up who may not already have a backhaul infrastructure in place through their AirSite® (AirSite is a registered trademark of AirNet Communications Corporation) Backhaul™ (Backhaul Free is a registered trademark of AirNet Communications Corporation) Free base station.

Press releases on the AirNet website suggest that the group have been having some recent successes in penetrating a range of markets with their products. In June 2004 they received their first North American large operator purchase order worth approximately \$3 million. The operator was undisclosed but was described as having approximately 1.7 million subscribers and serving rural and metropolitan service areas throughout the United States.

They also report sales to the Middle East and West Africa to help rapidly deploy developing GSM networks. These sales support the view that SDR could be of real benefit for quickly and inexpensively deploying wireless networks in developing countries where a substantial wired infrastructure doesn't exist. SDR then provides the opportunity for these networks to evolve and grow around demand.

Finally, they also report a \$1.4 million order from the US National Guard who is evaluating using the AirNet SDR base stations to rapidly deploy emergency response communications networks in response to major incidents. The use of SDR in the public safety sector is considered more in section 12.4.

12.2.2.3 Vanu Anywave™ Base Station

Vanu Inc is a US based company focused around the development of software radio products and began business in 1998 [322]. Recently they have received a lot of attention as in March 2005 their Anywave™ GSM base station became the first product to receive approval under the new US FCC software radio regulations. It was also awarded the 2005 GSM Award for Best Network Infrastructure [323]. The Anywave™ base station provides a software implementation of the BTS, BCS and TRAU units of the BSS and is run on a general purpose server. The base station uses COTS hardware components consisting of an antenna, wideband transceiver and general purpose processing platform (see Figure 12-3).

The Anywave™ (Anywave is a registered trademark of Vanu Inc.) base station currently supports GSM and can be upgraded to GPRS and EDGE. A CDMA upgrade is in development. The BSC connects to the backhaul network infrastructure via Gigabit Ethernet switches and is compatible with T1, microwave, fibre and satellite backhaul infrastructures.

The hardware software combination that received FCC approval included an ADC Digivance RF front end and was run on HP ProLiant servers. This product has already been deployed in rural Texas by Mid Tex Cellular who entered into a highly successful trial with Vanu in August 2003 [316].

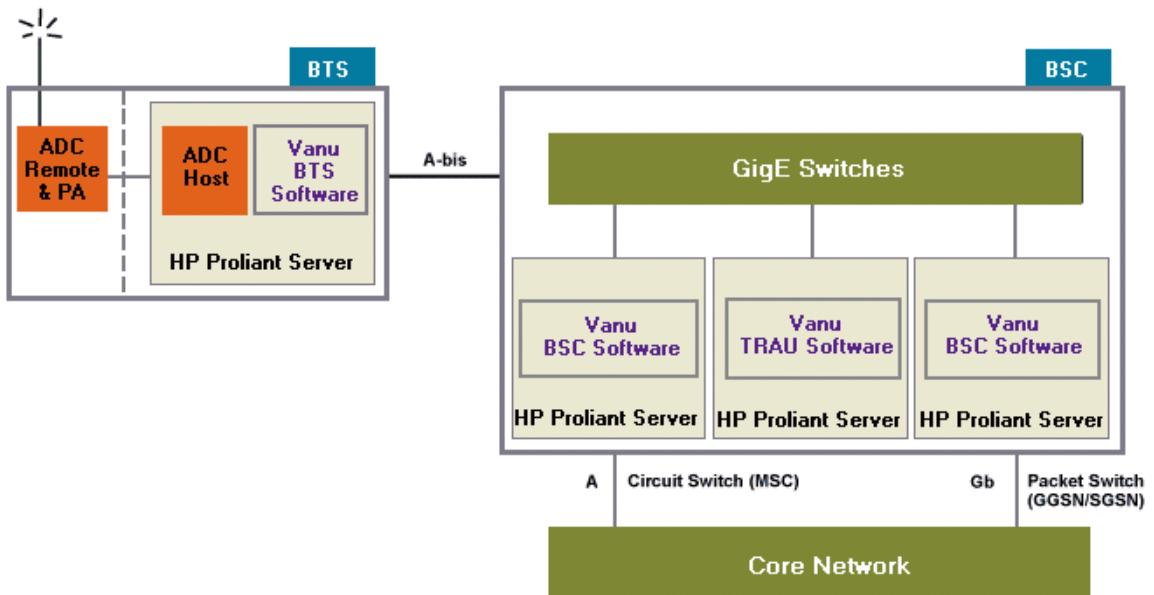


Figure 12-3: Vanu Anywave™ Base Station Architecture - Picture courtesy of Vanu Inc.

Mid Tex originally had a US TDMA network that they increasingly realised would need to be upgraded to the next generation of standard. However, they were unsure whether to upgrade to a GSM or CDMA solution and also were anxious about the costs involved. The trial with the Vanu base station successfully showed how they could concurrently run a TDMA and GSM network. It also gave them the ability to quickly and cheaply assess the markets readiness for the GSM service.

The trial also illustrated SDRs ability to remotely upgrade and fix bugs on the base station. As this was the first time the base station was deployed in a real scenario, problems emerged during the trial in areas such as call handover. However, as the team back in Vanu in Massachusetts came up with a solution a new software upgrade was made available to be downloaded to the base stations via a secure Internet link.

According to Mid Tex Cellular this solution has enabled them to significantly reduce operating costs and increase revenue due to the new GSM traffic [323]. Following this trial other operators such as AT&T and Nextel expressed interest in the Anywave™ base station demonstrating that the industry is interested in SDR developments [319].

Mid Tex Cellular's GSM system currently consists of 20 base station sites with close to 2000 subscribers. Roamers using the network include users from Cingular and AT&T. Vanu are in the process of also bringing GPRS to the network.

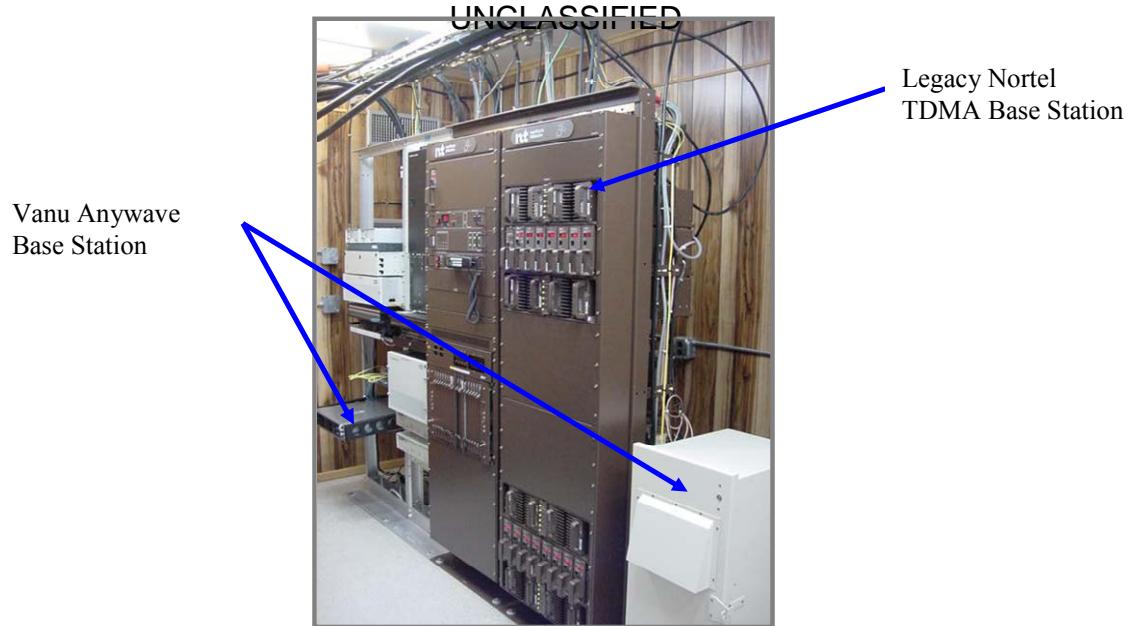


Figure 12-4: Vanu Anywave™ Base Station used in Mid Tex Cellular Trial - Picture courtesy of Vanu Inc.

12.2.2.4 Airspan AS.MAX Base Stations

In March 2005 Airspan released the first commercially available IEEE 802.16 or WiMAX base station [324]. WiMAX offers wireless broadband access to PDAs or laptops and is being discussed as a 4G standard by some.

The AS.MAX base stations uses picoChip's picoArray™ (picoArray is a registered trademark of picoChip Designs Ltd.) and a reference software implementation of the IEEE 802.16d standard. The picoArray™ is a flexible programmable platform that is ten times faster in processing power than any of today's leading DSPs [325]. picoChip also provide a full development environment and reference designs for many industry standards such as WiMAX, WCDMA and TDD.

This makes the picoArray™ a suitable platform for inclusion in SDR products and evolving standards like WiMAX. The AS.MAX base station promises to be upgradeable to the next generation mobile 802.16e standard and so offers a future-proof route to operators looking to begin rolling out WiMAX services.

Even though the WiMAX certification process hasn't yet begun, Airspan have already had a WiMAX network contract win [326] and AT&T are reviewing using this product [327].

As WiMAX is aimed at users of laptops and PDAs rather than mobile phone users this has been discussed in more detail in Section 12.3.1. However, it is worth noting the growing popularity of WiMAX, the ability of SDR to support it as an evolving standard and its potential to overlap with 3G and 4G as a high speed data service.

UNCLASSIFIED

12.2.3 3G Roll-out and SDR

In 2001 the 3GNewsroom was reporting SDR base stations as the key solution to 3G rollout problems [328][329]. The ability of SDR base stations to reconfigure on the fly and to support multiple protocols was thought to be the safest option for rolling out 3G with confusion over multiple 3G standards. Also SDR gave the opportunity to overlap 2G and 3G services and the ability to dynamically reconfigure the network between protocols depending on the call load. SDR also promised the ability to rapidly deploy trial networks to test the market's readiness. This gave a flexible solution that could support the migration of users to the new network and accommodate evolving standards.

However, looking at the situation in the UK in 2005, 3G is in the process of being rolled out [330] and SDR hasn't played the key role that was anticipated in 2001. SDR was considered for the 3G roll out but there were some key reasons against using the SDR branded base stations. These were:

- Lack of SDR UMTS solutions - None of the commercially available SDR base stations mentioned in Section 12.2.2, currently offer 3G solutions. It is feasible to combine GSM and UMTS in a single SDR device as experimental systems. However, none of these experimental systems have been commercially realised.
- Lack of Track History - Most network operators have a tried and trusted relationship with the major cellular equipment manufacturers like Siemens or Motorola. The major cellular equipment manufacturers also have wider product ranges and so can offer package prices to roll out the entire network rather than just the base station components (see Figure 12-2). Partnerships such as Vanu and HP show steps towards acquiring this track history.
- Lack of Compatibility with Existing Infrastructure – As SDR groups have mainly focused on base station solutions, they have the challenge of ensuring that their product will interface to equipment in the existing network such as the network sub-system and the operation sub-system. As the existing infrastructure has been supplied by one of the established cellular equipment manufacturers, the interfaces may be proprietary and so difficult to ensure compatibility with.

Importantly, this does not mean that the industry was not thinking about SDR solutions and in fact a review of the cellular industry does reveal a shift in thinking towards SDR techniques.

To first assess the industry's interest in SDR an industry map for the cellular mobile industry was drawn up and is shown in Table 12-1. This is by no means a comprehensive list of all stakeholders in the cellular industry worldwide but it does give an idea of some of the key companies and their attitudes. In order to assess the penetration of SDR into this industry the attitudes of each stage of the chain are examined.

UNCLASSIFIED

Semiconductor Houses		Manufacturers		Network Operators/ Service Providers		End Users
	SDR forum?		SDR Forum?		SDR Forum?	
Xilinx	Yes	Traditional		UK		Mobile phone users
picoChip	Yes	Lucent	No	Orange	Yes	
Texas Instruments	No	Ericsson	No	O2	No	
Analogue Devices	No	Motorola	Yes	Vodaphone	No	
Altera	Yes	Siemens	Yes	T-Mobile	No	
Qualcomm	Yes	Nokia	No			
Intel	Yes			US		
		SDR Start-ups		AT&T	No	
		AirNet	Yes	Cingular Wireless	Yes	
		Vanu	Yes	Nextel	No	
		WiMAX		Japan		
		Airspan	No	NTT DoCoMo	Yes	

Table 12-1: Industry map for Cellular Network Industry

UNCLASSIFIED

12.2.3.1 Semiconductor Houses

One key enabler for SDR is the availability of flexible, fast, programmable platforms such as FPGAs and DSPs. Generally, the main semiconductor houses have been quick to realise the Intellectual Property (IP) opportunities open to them through SDR. Many have already built reference designs for many of the cellular mobile standards. As can be seen in Table 12-1, many of them are members of the SDR Forum (an industry group established to promote the development of SDR techniques).

With FPGAs being one of the popular platforms for SDR, Xilinx naturally have a keen interest in it. They are members of the SDR Forum and have recently demonstrated the ability to run SCA compliant waveforms on one of their commercially available FPGAs [331].

Analogue Devices, although not members of the SDR Forum, have recognised the potential opportunities for their DSP and ADC products in SDR. Analogue Devices provided the ADCs and DACs in Vanu's Anywave™ Base station described in Section 12.2.2 [332]. Their TigerSHARC® (TigerSHARC is a registered trademark of Analogue Devices, Inc.) family of DSPs are also being promoted for implementing SDR in base stations [333].

There are also recent additions to the semiconductor group like picoChip. picoChip has developed their own fast, flexible, programmable platform in their picoArray™ product as discussed in Section 12.2.2. They have also built up IP by implementing a range of SDR building blocks targeted at the picoArray™ platform. Their implemented waveforms include W-CDMA and WiMAX.

12.2.3.2 Manufacturers

In 2002 PA Consulting were the first group to develop an all-software baseband system for WCDMA 3G base stations which ran on a DSP[334]. This was to be indicative of the trend, that although perhaps not branded as SDR products, FPGAs and DSPs were to become widely used in 3G base stations.

Due to commercial sensitivities, it is difficult to confirm how big a role SDR plays in the 3G base stations produced by the main cellular equipment manufacturers today. However, when the new High Speed Data variant of 3G, HSDPA, was introduced most manufacturers were able to upgrade already deployed base stations via a firmware software upgrade, indicating usage of reprogrammable devices.

It is also interesting to note that Siemens Communications Group have recently announced the use of Texas Instruments' TCI6482 1GHZ DSP in their new media gateway product that supports 2G, 2.5G and 3G mobile core networks. This is being promoted as a future-proof option due to its ability to be reprogrammed and support multiple waveforms [335].

Even the manufacturers who aren't part of the SDR Forum have indicated movement towards software based solutions. In Lucent Technologies' "Guide to GSM Network Migration" they describe the procedure for upgrading from GSM to GPRS as simple software only upgrade [336]. Also an article in Ericsson Review describing the architecture of their WCDMA 3G base station, described the extensive use of FPGAs and DSPs in the base band processing [337]. Although at first sight it seems that SDR has not made a big impact on the cellular network industry it actually has slowly but successfully been creeping in to the main manufacturers products, largely unnoticed.

As described in Section 12.2.2, SDR-branded products are available from AirNet and Vanu but face barriers such as interface issues, track history and lack of availability of 3G solutions. Also with the major manufacturers now using FPGAs and DSPs to make their products more flexible, these newcomers will have to work harder to prove what added benefits they can offer.

These additional benefits may include a wideband RF front end and the ability to reconfigure remotely and on the fly. These are key elements required for movement towards dynamic spectrum allocation schemes, spectrum trading and eventually cognitive radio. Intelligent antennas will also be key in order to support highly flexible network architectures that can support a variety of standards.

The issue of interfacing with existing cellular products may be made easier for SDR start-ups if initiatives like the Open Base Station Architecture Interface (OBSAI) are successful. This is an industry group aiming to standardise the key components of base stations and the interfaces between them. It splits base stations into four main blocks as shown in Figure 12-5. As SDR is focused around the base band processes having an agreed standard for the base band block would greatly assist manufacturers of SDR products looking to contribute to cellular network products.

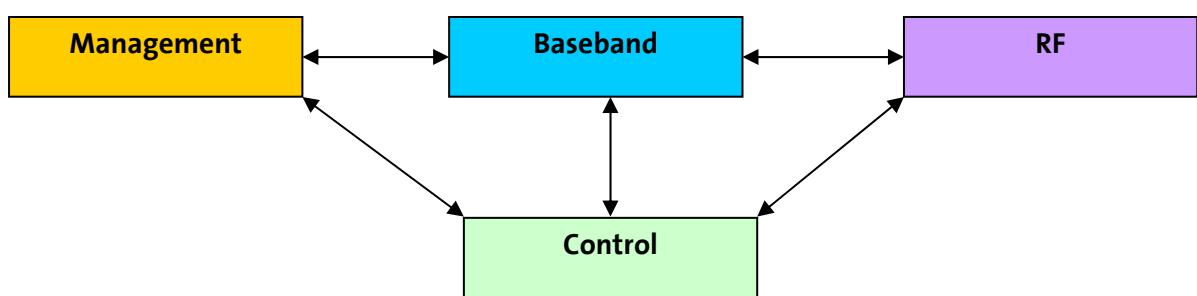


Figure 12-5: Main blocks of the OBSAI

12.2.3.3 Network Operators and Service Providers

The SDR Forum has been active in capturing the future requirements of network operators and their attitudes towards SDR. In June 2003 the SDR Forum R&D working group began the SDR Operators Market Requirements Survey (OMRS). As part of this, six of the SDR Forums' members who are network operators took part in a questionnaire. The main results of this were discussed in November 2003 in the SDR Forum's annual conference and are reported by Telephony Online as the following [338]:

- Half of the operators thought that, even after the introduction of 3G, 2G and 2.5G networks would continue to grow.
- The majority of operators said they plan to deploy 802.11, and half said they are planning to deploy 802.16
- The majority of operators saw a need for SDR handsets and half of them saw the potential for deploying SDR in their networks as high.

UNCLASSIFIED

- 4 out of the 6 operators reported the main benefits of SDR handsets as easier, fast bug fixes and air interface support. The other 2 operators thought that simultaneous multi mode operation was the main benefit.
- 5 out of the 6 operators said that they would be interested in a single base station capable of covering multiple bands.
- 3 out of the 6 operators said that they envision dynamic allocation of radio spectrum sometime in the future.

These results show a requirements set that is very much in line with the strengths of SDR. However, it is worth noting that those questioned were members of the SDR Forum and so in this way have already shown their interest in the development of SDR. The challenge for the SDR Forum is to extend this survey to network operators outside the forum in order to capture a more balanced view.

The reason network operators have been slow to support the new SDR product ranges may be because this type of change to their networks would be too abrupt. SDR may be seen as too much of an uncertain jump from existing cellular products. However, as described in Section 12.2.3, SDR has already been making its way into cellular network products from the main manufacturers. From this we can conclude that rather than being an abrupt change in the way base stations operate, SDR will be part of the gradual evolution of cellular equipment at a pace that network operators feel comfortable with.

Overall network operators should be in favour of SDR as it brings benefits to them such as the ability to offer multiple services and lower deployment and maintenance costs. Looking to the future the flexibility of SDR could also enable more efficient spectrum allocation schemes and spectrum trading that would result in the ability to support more users and gain increased revenues.

12.2.3.4 End Users

The end user should be very much in favour of SDR as overall it means that they will receive a greater variety of improved services quicker and cheaper than before. However, one of the key benefits of SDR is that it should allow multiple standards to coexist transparent to the user. The user will probably be unaware of what SDR is or its introduction. Their increasing demand for more services from a single handset will be a driver for the adoption of SDR into cellular products.

12.2.4 Mobile Phones

Chapter 11 has already raised some issues about the introduction of SDR to mobile handsets. It highlights that the requirements of mobile phone users are different to the network operators acquiring network infrastructure products. These include cost, weight, battery life, cosmetic design and number of additional features.

UNCLASSIFIED

The main issue with introducing SDR into handsets is that it requires the use of programmable devices which are generally power hungry and hence lead to reduced battery life and large devices. It is interesting to note that in Ofcom's report looking at the communications market in 2004, there were issues with the mass market acceptance of 3G handsets [330]. This was because the complexity of the standard meant that higher processing powers were required resulting in the handsets being expensive to produce. They were also large and heavy and didn't provide the required battery life. However, the report anticipates that these issues will be overcome in the next generation of 3G handsets.

The economic case for using SDR, as it is based around a standard platform, might not pay off in handsets. This is because the volumes are much greater and so can justify the initial outlay into the development of custom built ASICs.

SDRs ability to future-proof devices may not be of benefit to retailers of handsets as they want customers to frequently change their handset. With SDR, end users could still be charged for software upgrades to their mobile handset. However, they wouldn't get the corresponding cosmetic upgrade that comes with getting a completely new device.

SDR does provide the ability to support multiple waveforms on a single device and so ultimately could give an end user increased choice of services if incorporated into a mobile handset. SDR could also assist seamless roaming both at a national and international level. As processing platforms emerge that overcome power and size constraints, it is likely that SDR will make its way into handsets. This will probably be a gradual process as has been observed with the base stations.

12.2.5 Summary

Cellular networks are the most obvious and perhaps most lucrative market that SDR could penetrate. The benefits it could bring to this industry include a standard and therefore more economic hardware platform, future-proofing and easier bug-fixes through software upgrades and increased functionality and interoperability through the ability to support multiple standards. The flexibility of SDR could also move the industry towards more spectrally efficient allocation schemes, spectrum trading and even cognitive radio.

Examples of commercially available SDR base stations from AirNet, Vanu and Airspan have been reviewed and it seems that SDR is being recognised in the industry with Vanu's base station being the first to receive FCC accreditation as an SDR device.

Although it may seem that SDR has been slow to be taken up by the main cellular infrastructure manufacturers, a closer inspection has shown that programmable devices are key components of current 3G base stations. Perhaps the way forward for SDR in cellular base stations is not to be introduced as an entirely new product but instead as a natural evolution and upgrade to existing infrastructure. It could be argued that the most successful SDR products to date are the ones that have not been recognised as such.

12.3 SDR in other Commercial Applications

While the mobile cellular industry has so far been the main target for commercial SDR developments, there are other commercial application segments that should not be overlooked. If SDR start-ups continue to have difficulty penetrating the mobile cellular industry due to incompatibility with existing infrastructure, they may have more success supporting one or more of the emerging WLAN or WMAN standards.

These have the benefit of being new to the market and so could be rolled out more independently of existing mobile cellular networks. There is also the bonus that some of these standards are less complex to implement than 3G standards like UMTS and so might be more easily realisable in a software radio with current technology.

However, as each wireless standard has its own set of advantages it is unlikely that a single one will dominate so a device that incorporated all of them would be the ideal solution. SDR supports this vision.

This section looks at the potential additional features that SDR could bring to mobile devices. These may include high speed data links through WiMAX and WiFi or location services through Global Positioning System (GPS). With such a range of services on offer to the user, the form factor used to present them may well change from the traditional mobile phone. PDAs could prove to be the ideal platform as they are compact but can more easily display more information than a mobile phone.

The potential use of SDR in communications within the transportation industry is also briefly considered in this section. With traffic management and congestion charging becoming increasingly popular areas of interest, SDR may provide the ability to track vehicles across regional and national boundaries.

12.3.1 WiFi, Bluetooth and WiMax

As discussed in section 12.2, in the past the mobile cellular industry and fixed telephony industries have been kept separate. Traditionally the mobile phone user has been limited to voice and simple data services like SMS while the fixed user has been able to enjoy all of these plus high speed Internet access. However, the emergence of new wireless standards has meant that the fixed users are starting to enjoy their usual services with limited mobility. Suddenly, the fixed and wireless communications industries are being forced to overlap each other as the user base of both sectors start to merge.

Recently there seems to have been a flood of wireless standards entering the market such as Wi-Fi, WiMAX, UWB, Bluetooth, Zigbee and Near Field Communications (NFC). However, rather than being direct competitors to one another, they have each been developed with a specific application in mind.

NFC is being developed for very short range communications, in the region of a few centimetres [339]. It could be used, for example, to connect a mobile phone to a PC wirelessly just by holding the phone next to the PC. It could also be used in train stations to buy tickets from kiosks via your mobile phone. The key benefit of this standard is that that NFC device requires little power as it gets its power through RF coupling from the reader. This makes it ideal for use with mobile phones with Nokia, Samsung and Motorola all planning to incorporate it into their mobile phone products.

UNCLASSIFIED

Bluetooth and Zigbee fall into the category of personal area networks and are designed to provide wireless connectivity across a short range of up to about 10 metres. Bluetooth can support data rates up to 1Mbps and these standards are aimed at connecting PC peripherals wirelessly and have been used in wireless mobile phone headsets too. The benefit of these standards compared with wider range wireless standards is that they are relatively simple to implement and therefore cost less.

The wireless standard that has recently been responsible for giving high speed network connections to remote laptops users is the IEEE 802.11 family. These standards cover Wireless Local Area Networks (WLAN) and the popular IEEE 802.11b standard operates over ranges of up to 100m and supports data rates of 11 Mbps. The emergence of so called WiFi hotspots providing WLAN access points for in range WLAN users provide the opportunity for laptop and PDA users to connect to a network wirelessly in a variety of public places. The deployment of hotspots by operators in the UK at least is set to rapidly grow as reported by Ofcom's 2004 assessment of the communications market [330].

As one of the key offerings of 3G is high speed data access, WiFi could be seen as a competing standard. A recent report by Strategic Analytics reportedly indicates that the introduction of WiFi will have a significant impact on 3G profits [340]. However, it would seem that rather than competing against this new wireless standard, cellular operators have decided to embrace it. In the UK it is T-Mobile and BT Openzone who are leading the roll out of WiFi hotspots [330].

At the next level of range the IEEE 802.16 or WiMAX standards are being developed. These provide long distance wireless broadband connectivity at ranges of up to 48km and data rates of up to 75 Mbps. While the fixed version of this standard has been set, the mobile equivalent is still under development and potentially several years from being realised [340]. Once this mobile potential is developed WiMAX could again be seen as a competitor to 3G [341]. However, crucially WiMAX has been designed with high speed data transfer in mind and so is not suited to voice traffic. Again the cellular industry's attitude has been to embrace this new standard with cellular equipment manufacturers such as Nokia, Ericsson and Motorola taking part in the industry group the WiMAX forum [342].

WiMAX is also being thought of as a technology for linking multiple WiFi hotspots together. Intel recently demonstrated the UK's first wireless network based around WiMAX [343]. It was given a special licence from Ofcom to carry out research and testing of its WiMAX network based at its Swindon headquarters. The trials of the Intel's WiMAX service were centred on providing wireless broadband to a Science Museum 6km away that suffered from a lack of wired network infrastructure. The team set-up and used the WiMAX connection to link WiFi networks in each of the museum's storage facilities to provide a central electronic inventory.

Each of these wireless standards have their own set of benefits that distinguish them from one another. The main cellular manufacturers and operators have been interested in incorporating these into their current product portfolios to provide the best quality, economy and variety of service possible to the end user. With so many standards the end user could be easily be confused and so the ultimate goal is to transparently incorporate these services into one device. The ability of SDR to combine multiple standards and support the development of cognitive devices that can adapt with their environment fits this vision perfectly.

UNCLASSIFIED

BT has been leading the way in terms of developing a product that merges the fixed and wireless markets. Their research project BT Bluephone aims to develop a mobile phone that will normally operate as a GSM phone [330]. Once within the user's home it will make a Bluetooth connection to the fixed household BT broadband connection and provide the user with broadband connectivity and land line priced voice calls. This provides the user with the benefits of both services but with the luxury of any switch over being transparent to them.

In terms of commercially available SDR products that support these more recent wireless standards, AirNet's SDR base station mentioned in Section 12.2.2 can be configured to operate as a GSM, GPRS, EDGE or WiMAX base station. Airspan described in 12.2.2 have released an SDR WiMAX base station that will be involved in trials with service providers in 2005 [344].

12.3.2 GPS

Another benefit of SDR is that it could help to incorporate features that are of great use on a mobile device but have traditionally been kept in separate chipsets. GPS is one such technology that falls into this category. GPS is useful to mobile devices for a number of reasons. From a technical point of view it can aid synchronisation as we move towards more packet based systems [345]. In terms of functionality, users could benefit from navigation aids in their mobile phone. When 3G was evolving it was envisaged that the increased data rates of 3G combined with GPS capability in mobiles could be used to give location based services to 3G users such as directions to their nearest train station. Companies like Sepura have been incorporating GPS into their TETRA handsets [346]. This is of benefit to public safety users as it means central controllers can quickly and reliably locate their team, check for problems and decide how to deploy them most efficiently. Such capabilities are also of use to the highway agencies.



Figure 12-6: Sepura SRP2000 sGPS TETRA Handportable - Picture courtesy of Sepura Ltd

UNCLASSIFIED

SDR provides the opportunity to reduce costs on mobile devices incorporating GPS by including a software GPS implementation in the SDR rather than having to add a separate GPS chip. An additional benefit is that as GPS standards change, an SDR implementation could be upgraded with a simple software download rather than a hardware change.

SDR implementations of GPS are feasible with current technology as proven by Navsys' software GPS system testbed [347]. While the testbed is an experimental system and so would need to be reduced in size to be incorporated into handsets, it does prove the concept.

12.3.3 PDAs

If SDR gives the user the ability to access voice, data and location services the traditional mobile phone handset might not be the best form factor to present this information to the user on. This may mean a movement towards a wider mixture of mobile devices such as laptops, mobile phones and PDAs. The PDA may in fact be targeted as the first form factor of such multi functional devices as they provide a good balance between mobility and functionality. Also being based around a general purpose processor PDAs could provide the ideal hardware for running platform independent SDR modules.

Recently, Vanu and General Dynamics Decision Systems have demonstrated a handheld software radio prototype [322]. The prototype consists of a General Dynamics radio transceiver and an iPAQ containing a 206MHz StrongARM processor which runs the Vanu software radio under Linux. All hardware blocks used in the prototype are COTS products. The current prototype can support an analogue, two-way FM radio and APCO Project 25, making it applicable to public safety users.



Figure 12-7 Vanu Handheld Software Radio Prototype - Pictures courtesy of Vanu Inc.

UNCLASSIFIED

While this is not a commercial product it is encouraging that the prototype, demonstrating two waveforms, was built in a six week period. This shows how easily the flexibility of existing SDR products can be transported and exploited if the RF front-end and processing platform are available to support it. This prototype shows the power of Vanu's philosophy of building software radios that are completely platform independent and can be ported to any COTS processor. If the rest of the SDR community follows this role SDR may boost the popularity of PDAs.

12.3.4 Transportation: Automotive and Commercial

Traffic monitoring and congestion control have become key issues worldwide. Nations are increasingly looking to improve the efficiency of their transport infrastructures through Intelligent Transport Systems (ITS). Following recommendations from Automotive Innovation and Growth Team, an ITS centre of excellence has been recently established in the UK.

The main aim of ITS is to add information technology to existing transport infrastructures and vehicles in order to collate information about real time traffic levels and to use this information to use the transport infrastructure as efficiently as possible. One such application is in congestion charging. In order to ensure that cars within a city's limits were paying the appropriate congestion charge, it would be useful to be able to track vehicles. ITS could assist with this. In terms of commercial transport, it would be useful for hauliers to be able to track their vehicles progress and plan efficient deployment. ITS is particularly relevant to the UK at the moment with pay-as-you-go road charging being discussed by the transport secretary as a future technique for tackling congestion [348].

There are a number of reasons why SDR could be incorporated into ITS. Being able to track vehicles implies radio communications between the vehicle and central control centre. However, as a vehicle moves across regional and national borders the same ITS communications standards may not be supported. SDR could provide the opportunity for vehicles to reconfigure and seamlessly roam and remain in contact as they crossed wireless standards boundaries.

As previously mentioned, SDR also brings the opportunity to incorporate additional features to mobile devices such as GPS. The incorporation of GPS and location based services would be very useful for ITS.

The third benefit is the ability of SDR to remotely upgrade devices. If ITS radio devices are fitted to vehicles they will be very difficult to recall and upgrade as standards are enhanced. This could lead to a progress gap between new and old vehicles. However, as SDR offers the opportunity to remotely upgrade devices; this could help overcome this issue.

ITS is still quite an immature market with many industry focus groups being quite new and just starting research and development projects to develop the technology to support ITS concepts. From this point of view, it is an ideal fresh new market for SDR to dominate from the start without the issues of established markets like the cellular industry. However, it may take some time for ITS products to become commercially available.

UNCLASSIFIED

12.3.5 Summary

It is important to remember that while the cellular network industry has been the main focus of commercial SDR developments, there are other key applications that should not be neglected.

Recently there has been a flood of wireless standards emerging such as WiFi, WiMAX and NFC as well as some more familiar ones like Bluetooth. This group of wireless network standards are generally aimed at giving fixed users increased mobility but maintaining high data rate services. On the other side, mobile users are starting to acquire higher speed data services and so an overlap is emerging between the fixed and wired telecommunications industries. However, this does not necessarily mean increased competition and indeed many of the cellular operators and equipment manufacturers are embracing the new standards and adding them to their product portfolio. BT's Bluephone project is an excellent example of this.

As each standard has its own advantages, there should be a place for all of them in future wireless devices. However, there is a strong risk of confusing the consumer. SDR is the ideal solution to this problem as it provides the opportunity to support and roam between multiple wireless standards. Current technology may only permit the combination of a subset of standards but as SDR develops and becomes more wideband this goal will become more feasible.

As well as incorporating additional wireless standards into a mobile device, SDR also presents the opportunity to incorporate extra features. GPS is a good example of this. GPS can provide the mobile user with better synchronisation and also location based services. Normally, GPS is included as a separate chip on mobile handsets but a software GPS receiver has been demonstrated showing that it could be incorporated within an SDR.

With SDR potentially bringing so many additional services to mobile users the form factor that they are presented in may well change from the mobile phone. A PDA would provide a good balance between mobility and functionality given the additional high speed data and location based services on offer. Interestingly, a PDA has also been recently used by Vanu as a platform for its prototype handheld software radio.

Traffic control and congestion is an area that is increasingly an issue worldwide. Many nations are discussing ITSs that bring information technology to help monitor and manage transport infrastructures. The ability of SDR to incorporate GPS, provide seamless roaming across national boundaries and to provide remote upgrades makes it attractive to this industry. Also, as ITS is a relatively new area it is a market that SDR could dominate from the outset. However, due to the markets immaturity it may be sometime before products are available.

12.4 SDR in Public Safety Mobile Radios

The JTRS programme shows how the US defence sector has recognised the ability of SDR to promote interoperability between forces. Traditionally, the wireless communications networks of each of the armed forces have been treated in isolation, with each group adopting their own wireless standards. This means that in joint operations the army are unable to directly communicate with the navy or the air force. The problem is worsened in international operations.

Given the success of SDR in diffusing interoperability issues in the defence industry, the next logical market to examine where similar situations could arise is the public safety sector. As in the defence sector, scenarios can arise, such as the attacks on the World Trade Centre, where multiple groups of first responders, such as police, ambulance crews and fire fighters, all need to work together as a team and therefore communicate.

This section examines the interoperability issues that plague the public safety sector and looks at how SDR might be able to help. The requirements of the public safety sector are examined as there are some key differences between this and the commercial cellular industry described earlier. As the UK has recently been quite active in harmonising public safety sector communications developments in this market are described. This section finishes by discussing emerging SDR products targeted at the public safety sector and attitudes towards SDR within this market.

12.4.1 The Need for Harmonisation across the Emergency Services

Lack of interoperability across the public safety sector is a widely recognised issue. This issue is actively being debated in the US with multiple public safety working groups and programmes emerging. For example, the US Department of Homeland Security's SAFECOM programme is an umbrella programme with the ambition of improving public safety response through more effective and efficient interoperable communications [349]. The National Institute of Justice also have a complimentary programme called CommTech which focuses more on evaluating the technology developments that could assist interoperability.

There are also working groups that provide a focus for key industry stakeholders to debate the issue of interoperability. One such group is the National Task Force on Interoperability who, in their brochure, has highlighted how lack of interoperability has still been an issue during recent crisis situations. They highlight how in the Oklahoma bombing in 1995 and more recently in the September 11th attacks in 2001, different groups of first responders had difficulty communicating directly with each other due to differences in radios [350].

While the US is working hard to tackle this issue, it seems that Europe may be ahead of them with their TErrestrial Trunked RAdio (TETRA) Private Mobile Radio (PMR) standard. TETRA is the only PMR standard that has been developed by ETSI and is used in a variety of sectors including public safety, utility, transport and defence. Although it is supported by many European countries, there are still competing standards to consider like the French TETRAPOL system.

UNCLASSIFIED

One of the key reasons for lack of interoperability within the public safety sector is diversity and age of radio equipment across regional and national boundaries. This is a result of independent procurement of radio equipment across the emergency services [351] and is similar to the scenario experienced in the defence sector where years of independent procurement of communication systems by the different forces have resulted in lack of interoperability between them all [352].

As the defence sector has embraced SDR as the key to unlocking this interoperability issue, there is also an interest in bringing SDR to the public safety sector. SDR is being noted as a key technology in this area. On the US National Public Safety Telecommunications Council (NPSTC) website they cite SDR as one of the key issues that they are currently examining. The SDR Forum have also recognised their potential impact on this sector and in April 2004 started a public safety special interest group [353].

12.4.2 Requirements of Public Safety Mobile Networks in Contrast to Commercial Cellular Networks

A recent study by Venture Development Corporation (VDC) showed that 88% of those from the US public safety sector who were included in their study indicated SDR technology could help their interoperability issues [354]. However, given this demand the public safety sector hasn't yet seen a commercially available SDR system.

One of the key reasons for this is that the commercial efforts of the SDR community have been focused on developing products for the cellular network industry as described in Section 12.2. While the emergency services use public mobile cellular systems for a lot of their day to day communications, this doesn't replace the need for PMR networks.

The public safety sector has a different set of requirements to the average mobile phone user such as security, availability, low delay and reliability. For these reasons, the public safety sector can't rely solely on public mobile cellular systems.

This was demonstrated during the response to the September 11th attacks [355]. The PSTN became completely overwhelmed and as public mobile cellular systems rely on PSTN backhaul connections public mobile phones became unusable. Fortunately, in this case it was planned in advance that the cellular network providers would respond to such a situation by deploying mobile cellular on wheels base stations to rapidly deploy a cellular network that was independent of the PSTN and could support the public safety services.

This requirement for reliable radios that offer high availability means that the public safety sector generally use old, tried and tested equipment rather than supporting more risky cutting edge technology that has yet to be proven. The dependence of SDR on software adds an additional concern over the security of the device and raises the question of how easily it could be tampered with.

While aesthetics is key in public cellular equipment, it is not so much of a concern to the public safety sector. Public safety radios need to be mobile but size is not as much of a concern as it might be in the public cellular industry. Emergency services handsets also need to operate in quite rugged conditions and will usually experience high wear and tear on a daily basis. For this reason, handsets are replaced on quite frequent basis and so the ability of SDR to future-proof the handset in particular might not have as much relevance in this sector.

UNCLASSIFIED

Interoperability is a key requirement of the public safety sector. The flexibility of SDR and its ability to support multiple waveforms is therefore a key benefit. It is also worth noting that the traffic levels of the emergency services has distinct peaks and troughs and so there may be the opportunity for them to generate revenue through spectrum sharing schemes which the flexibility of SDR could help to facilitate. However, the flipside of this is that the emergency services require guaranteed use of the spectrum when needed. This means that they may be opposed to developments like cognitive radio.

12.4.3 Public Safety Market within the UK

The issue of interoperability in the public safety sector is very much worse in the US than it is within the UK. The police services in the UK have recently upgraded their PMR equipment to a TETRA solution. Procurement of this nationwide system has been led by the Police Information Technology Organisation (PITO) [356]. The network selected by PITO was O2's Airwave system which provides a secure digital voice and data service based on the TETRA standard [357]. As well as being a digital radio it can also be used as a mobile phone or data terminal. The roll out of Airwave in police forces across the UK was completed in May 2005.

The key advantages of Airwave are that it provides interoperability amongst UK police forces across regional boundaries. If adopted by the ambulance and fire services, this would further extend its interoperability. Ofcom has also recently approved Electricity Networks Association (ENA), London Underground and Heathrow Express as users of the Airwave service [358].

The O2 Airwave network also benefits from being part of the critical national infrastructure and so is designed to provide a high availability service that should not be disrupted by a national crisis.

The UK fire services are currently deciding on a new PMR network to use. Funding of this programme comes centrally from the Office of the Deputy Prime Minister under the Firelink programme [359]. In partnership with Marconi, the O2 Airwave system is again a serious contender for this programme as it would give the fire services the same PMR system as the police. However, as Airwave is a TETRA system which is an open European standard it should be possible for other TETRA based equipment to be interoperable with it. This would include PMR equipment not just across UK services but also across European borders.

Even though the UK seems to be tackling the issue of interoperability, SDR could still offer benefits to this market. SDR does offer the ability to future-proof radios to ensure that if the TETRA standard is upgraded and improved, the UK emergency service can avail of it. Rival standards to TETRA such as TETRAPOL do exist, so SDR's ability to support multiple waveforms could again be of benefit.

12.4.4 Emerging SDR Products within Public Safety

While there have not been any commercially available SDR public safety radios, there have been some signs of SDR making its way into this sector.

UNCLASSIFIED

In October 2004 AirNet, as mentioned previously, announced that they had received a \$1.4 million order from the National Guard for their RapidCell™ (RapidCell is a registered trademark of AirNet Communications Corporation) SDR base station to be used in conjunction with AirSite® Backhaul Free™ units [360]. The units will be evaluated by the National Guard as a technique for ensuring emergency communications remain reliable even in the worst of circumstances. The ability of AirNet's product to provide a high capacity, cost efficient, rapidly deployable broadband macro cell makes it very appealing to incident response groups.

As mentioned in section 12.3.3, Vanu and General Dynamics Decision Systems have demonstrated a handheld software radio prototype. Vanu have published a "notional handheld software radio product concept" describing how this prototype could, perhaps be further developed into a commercial product [361]. The product concept focused on for this handheld software radio is the "Vanu Universal Public Safety Radio", showing the group's acknowledgement of public safety as a key area for software radio. The key benefit promoted for this concept product is its ability to overcome interoperability issues by supporting multiple public safety waveforms.

12.4.5 Summary

Traditionally defence communication systems have been procured in isolation which led to issues of interoperability between forces and nations. In the US, SDR is being used as a technique for overcoming these interoperability issues. The public safety services are in a similar situation to the defence sector with different emergency services procuring communication systems in isolation. It has been shown that interoperability is a key concern in the public safety sector. This is most apparent in the US and has been highlighted in reflections of recent incidents, most notably the attacks on the World Trade Centre.

SDR offers the ability to overcome these interoperability issues in the public safety sector in a similar way to the JTRS programme in the defence industry. An examination of the requirements of this sector has shown that SDR could have additional benefits to offer this sector on top of interoperability. End users require a high quality, reliable service that can be rapidly deployed and will remain operational whatever the circumstances. Commercial cellular networks rely on the PSTN and so become unreliable in times of emergency. There are cellular on wheels solutions for overcoming this link with the PSTN. However, SDR provides the opportunity to have flexible, rapidly deployable base stations that can be upgraded and adapted depending on the users it needs to support.

In Europe, public safety communications interoperability is not as much of an issue as in the US with standards like TETRA emerging. In the UK in particular, the police services have just completed roll-out of their Airwave TETRA compliant system with the fire services also looking to upgrade soon too. However, competing standards like TETRAPOL and possible upgrades to TETRA mean that SDR might still have a role to play in the European public safety sector.

UNCLASSIFIED

There has been recognition of the connection between SDR and public safety with the start-up of groups like the SDR Forum's public safety special interest group. Industry relationships are also starting to form as illustrated by the evaluation of the US National Guard of AirNet's SDR base stations. Also, the concept of Vanu developing an universal public safety radio from their existing handheld software radio prototype shows further proof that the SDR community, though perhaps traditionally more focused on penetrating the commercial mobile cellular industry, has also made progress within the public safety sector.

12.5 Forecast of SDR Roll Out

Having reviewed the benefits that SDR can bring to a variety of markets with today's technology, it is useful to reflect on the advantages that SDR could bring in the future. This will highlight the areas where the further development of SDR will continue to reap benefits long into the future.

The second part of this section compares and contrasts the popularity and potential benefits of SDR across the main target markets reviewed in this report. This will help build up a forecast of SDR developments and deployment over the coming years and highlight the key barriers to SDR's success in the commercial world.

12.5.1 Future Benefits of SDR

So far this report has concentrated on reviewing the benefits of SDR that either are or shortly will be within reach of today's technology and has referred to a number of commercially available SDR products that are delivering these benefits today. However, it is also worthwhile looking at where SDR will fit into the long-term future of communications.

One technology that is being heralded as the future of communications is multiple-input, multiple output (MIMO) devices. MIMO devices use multiple transmit and receive antennas to exploit the multi-path characteristics of wireless channels and achieve higher capacity wireless links. This technique is the subject of a lot of development work and is predicted by many to have a big impact on communications devices of the future. However, this technique is still a relatively long-term research area with hurdles such as incorporating multiple antennas into a handset still to be overcome [362].

As discussed in Chapter 5, if MIMO will be used in future standards, it is important that SDR supports it for SDR to have a long-term future. The link between MIMO and SDR is well illustrated by QinetiQ's JASMINE Wideband Multi Frequency MIMO Channel Sounder shown in Figure 12-8 [363]. The complexity of working with an eight element antenna array meant that the baseband processing needed to be built around a software radio type implementation. This has also given the group the flexibility to implement alternative sounding waveforms as required.

Another key area for the future direction of the communications industry is obtaining more efficient usage of the radio spectrum and the movement towards spectrum trading. While spectrum trading may seem a far off concept, Ofcom have recently announced that by 2010 they are aiming to have 71.5% of the spectrum being allocated under "Market Forces" [364].

UNCLASSIFIED

The ability of SDR to support novel techniques for dynamic spectrum allocation has already been shown in Chapter 9 (and is considered in more detail in Chapter 13). Cognitive Radio was also discussed as a key enabler of more complex dynamic spectrum allocation schemes. As Cognitive Radios are devices that have an awareness of their environment and can adapt to improve their performance based on this information, they are ideally suited to locate and make use of unused pockets of spectrum. SDR supports Cognitive Radio in that in order to adapt its performance the Cognitive Radio must be flexible.



Figure 12-8: QinetiQ's JASMINE Wideband Multi-frequency MIMO channel sounder

The conclusions of this work were used to build up a dependency chain showing the link between SDR, Cognitive Radio and Dynamic Spectrum Allocation and this is shown in Figure 12-9. This relationship means that developments in SDR will drive developments in Cognitive Radio which in turn will drive developments in dynamic spectrum allocation. If spectrum trading and dynamic spectrum allocation are the future of the communications industry, SDR will almost certainly be part of it too.

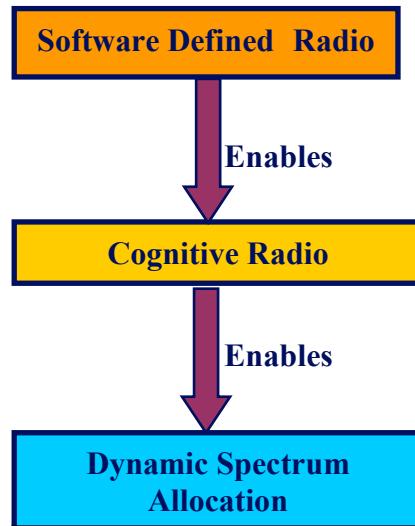


Figure 12-9: Dependency Chain

12.5.2 Comparison of SDR Popularity across Market Sectors

Table 12-2 and Table 12-3 show a comparison of the key benefits SDR across the markets reviewed in this report. In addition, Table 12-4 provides a similar comparison for the key issues. These three tables summarise the findings of the more detailed analysis of each market that has been described in the previous sections.

For each market, each of the benefits and issues of SDR are assessed and given a ranking according to their relevance to that market:

- “Key” describes a benefit that is highly important to exploit in order to gain success in the appropriate market.
- “Desirable” describes a benefit that would be very useful but lower priority to the “Key” benefits.
- “Mixed” describes benefits that particular market may view both as a benefit and a concern.
- “Against” describes a potential SDR benefit that is not of use to that particular market.

Similarly, issues with SDR deployment are also described across the market sectors and ranked.

The transport sector offers an interesting new market that SDR could potentially dominate from the start. While it is still relatively immature it would be worthwhile for SDR interest groups to keep involved with ITS initiatives and research. The public safety sector again is another market that SDR could dominate, particularly in the US. If the JTRS programme is highly successful a similar SDR public safety programme could be the obvious follow on.

UNCLASSIFIED

Benefits	Cellular Networks	WiFi, WiMAX, Bluetooth	GPS	Transportation	Public Safety
Decreased development and hardware platform costs	Key Cost is main factor driving cellular sales as service is similar across operators.	Desirable Cost is always a concern but as the new wireless standards give users new services they should be willing to pay for additional functionality.	Key SDR will only aid GPS if it can be more efficiently implemented in mobile devices than it currently is.	Key If an ITS-type device is to be fitted in all vehicles the costs could be very high.	Desirable Cost is a factor but reliability and interoperability will be the most significant issues.
Faster time to market	Desirable Most operators already have an established position and are aware of new standards emerging.	Key The ability to quickly incorporate these new services into their old portfolio could give them a commercial advantage.	Desirable GPS is already on the market so this isn't so critical.	Key ITS is a new and emerging market so being the first to release could give a manufacturer an advantage.	Desirable Not as critical as procuring a new public safety system is a long process.
Easily fix post manufacture bugs	Key Remote bug fixes would greatly reduce costs.	Key Probably not as essential as in the cellular network industry but still would reduce costs.	Key Remote bug fixes would greatly reduce costs.	Key Once ITS devices have been deployed, particularly on vehicles, they will be difficult to recall.	Key Remote bug fixes would greatly reduce costs.
Smooth upgrade and migration path	Desirable The ability of SDR to pre-empt all future standards is debatable.	Key Wireless standards are evolving so keeping pace with the latest is important.	Key Changes to the ranging codes used in GPS are being discussed.	Desirable As ITS is new, standards will probably evolve and change quite rapidly.	Desirable It is key that all services have the same compatible, up-to-date standards implemented. The turnover of public safety handsets is quite high so are frequently replaced.

Table 12-2: Comparison of SDR Benefits Across Markets - Part I

UNCLASSIFIED

Benefits	Cellular Networks	WiFi, WiMAX, Bluetooth	GPS	Transportation	Public Safety
Support multiple waveforms	Key (particularly in the US) US have multiple 2G standards to support. Worldwide there is a desire to support not just GSM or 3G but also WiMAX and WiFi	Key Rather than competing with cellular mobile services, devices that can support the most appropriate standard for the situation will give the user the best overall service.	Key The ability to merge GPS into a single SDR platform is key to making GPS more widely available.	Key The ability to support multiple standards across national borders is key for ITS.	Key Ensuring interoperability across borders and services is a high priority.
Supports cognitive radio	Mixed Operators will be in favour of fitting more users into their allocated spectrum but cautious about opening up their spectrum allocation to others.	Key One of the key issues with WiFi is interference in the ISM bands. Cognitive devices could help use this band more intelligently.	Against Due to the nature of this service it is unlikely that it will be possible to share GPS bands.	Desirable Cognitive radio could ease the problem of ITS having to negotiate a spectrum licence.	Mixed The spectrum usage of public safety services is peaky so there is the opportunity to trade. However, they need guaranteed QOS level and availability when they do need it.
Compliments MIMO	Desirable If MIMO is the future then eventually it would be advantageous for SDR to support it.	Desirable If MIMO is the future then eventually it would be advantageous for SDR to support it.	Desirable If MIMO is the future then eventually it would be advantageous for SDR to support it.	Desirable If MIMO is the future then eventually it would be advantageous for SDR to support it.	Desirable If MIMO is the future then eventually it would be advantageous for SDR to support it.

Table 12-3: Comparison of SDR Benefits Across Markets - Part II

UNCLASSIFIED

Issues	Cellular Networks	WiFi, WiMAX, Bluetooth	GPS	Transportation	Public Safety
Manufacturer Track Record	<p>Key Concern</p> <p>Cellular operators have existing relationships with manufacturers and will be hesitant to change these without good reason.</p>	<p>Moderate Concern</p> <p>As these standards are new, manufacturers have not yet established leading positions in this particular market.</p>	<p>Key Concern</p> <p>There are already GPS products that work well so an SDR equivalent would have to prove itself.</p>	<p>Not a Concern</p> <p>As ITS is a new market there hasn't been time to establish a track record.</p>	<p>Key Concern</p> <p>The public safety market needs a solution they have confidence in and will avoid new and hence high risk technologies.</p>
Interoperability with existing infrastructure	<p>Key Concern</p> <p>Compatibility with existing backhaul and control infrastructures is a key requirement.</p>	<p>Moderate Concern</p> <p>The ability to overlap with cellular services is desirable but these new standards aren't necessarily tied to an existing infrastructure.</p>	<p>Not a concern</p> <p>There is no backhaul infrastructure to link with.</p>	<p>Not a Concern</p> <p>As ITS is a new market, the infrastructure is yet to be defined.</p>	<p>Moderate Concern</p> <p>The ability to link into existing backhaul infrastructures is desirable.</p>
Size of handset	<p>Key Concern</p> <p>Aesthetics is key in the mobile phone industry.</p>	<p>Moderate Concern</p> <p>Users of the new wireless standards are generally laptop or PDA users and are used to larger devices but would prefer increased mobility.</p>	<p>Key Concern</p> <p>GPS can already be incorporated as an additional chip so an SDR implementation would need to beat this.</p>	<p>Moderate Concern</p> <p>If ITS devices are to be fitted to a vehicle, they need to be mobile but not necessarily as small as mobile phone users would demand.</p>	<p>Moderate Concern</p> <p>The handset must be mobile and easy to use but aesthetics isn't as big a concern as with mobile phone users.</p>

Table 12-4: Comparison of SDR Issues Across Markets

12.5.3 Predicted Trends for SDR Deployment

The commercial and technical barriers make it very difficult to predict how, where and when SDR will be rolled out by the communications industry. An article from January 2000 by the SDR Forum's Marketing Group reported an anticipated annual global demand in the commercial wireless market by 2005 for 130 million SDR-enabled appliances [317]. In August 2004 Venture Development Corporation released a white paper analysing the demand for SDR in North America and Europe. It forecasts growth in SDR device revenues to grow from \$1.1 billion in 2003 to \$5.3 billion in 2007 [365]. This shows how a large take up of SDR has been anticipated for the past five years.

In recent years the defence industry has been the main sector to embrace SDR with the JTRS programme giving it a real boost. It has been shown that SDR has already been making its way into the cellular industry, both as SDR branded products and the introduction of reconfigurable devices to main stream base stations.

Due to the technical and commercial issues (as exemplified by Table 12-2, Table 12-3 and Table 12-4) that still surround SDR, it is difficult to forecast its future deployment. The deployment of SDR has happened but the speed with which it is adopted will depend on the success of initiatives like OBSAI, partnerships between SDR start-ups and main manufacturers and future regulation of SDR devices.

Figure 12-10 gives an overview of the areas identified by this report where SDR is likely to have an impact on in the near future. Due to technical limitations, base stations rather than handsets have been the first to benefit from these. As SDR matures, it should play a part in the smooth migration from 3G to 4G with base stations supporting both standards.

As size and battery limitations are overcome, SDR will be incorporated in mobile devices. At first mobile SDR devices will be larger wireless devices such as laptops and PDAs, where high power consumption isn't as much of a concern and eventually move towards handsets. This report has already discussed projects that are underway to incorporate multiple waveforms into handsets and this will probably be the key reason for bringing SDR to handsets.

As SDR comes to both base station and handsets, more intelligent spectrum allocation schemes may emerge. With targets like Ofcom's ambition to bring market forces into spectrum allocation by 2010 this could in turn lead to more dynamic spectrum allocation schemes being implemented and spectrum trading. The flexibility of SDR will very much aid this transition.

The transportation industry holds good opportunities for SDR developers but will need time to scope and define the visions of ITS first. In the public safety market the issue of interoperability, particularly in the US, urgently requires resolution. However, procuring such systems take time and the public safety industry will need convincing of the reliability of SDR technology before they choose to embrace it.

If the JTRS programme is highly successful a similar programme based around public safety communications may be established. The public safety sector could also benefit from releasing part of their spectrum allocation through spectrum trading. It will however, be some time before this happens as they will first need to be convinced by the success of such initiatives in other sectors.

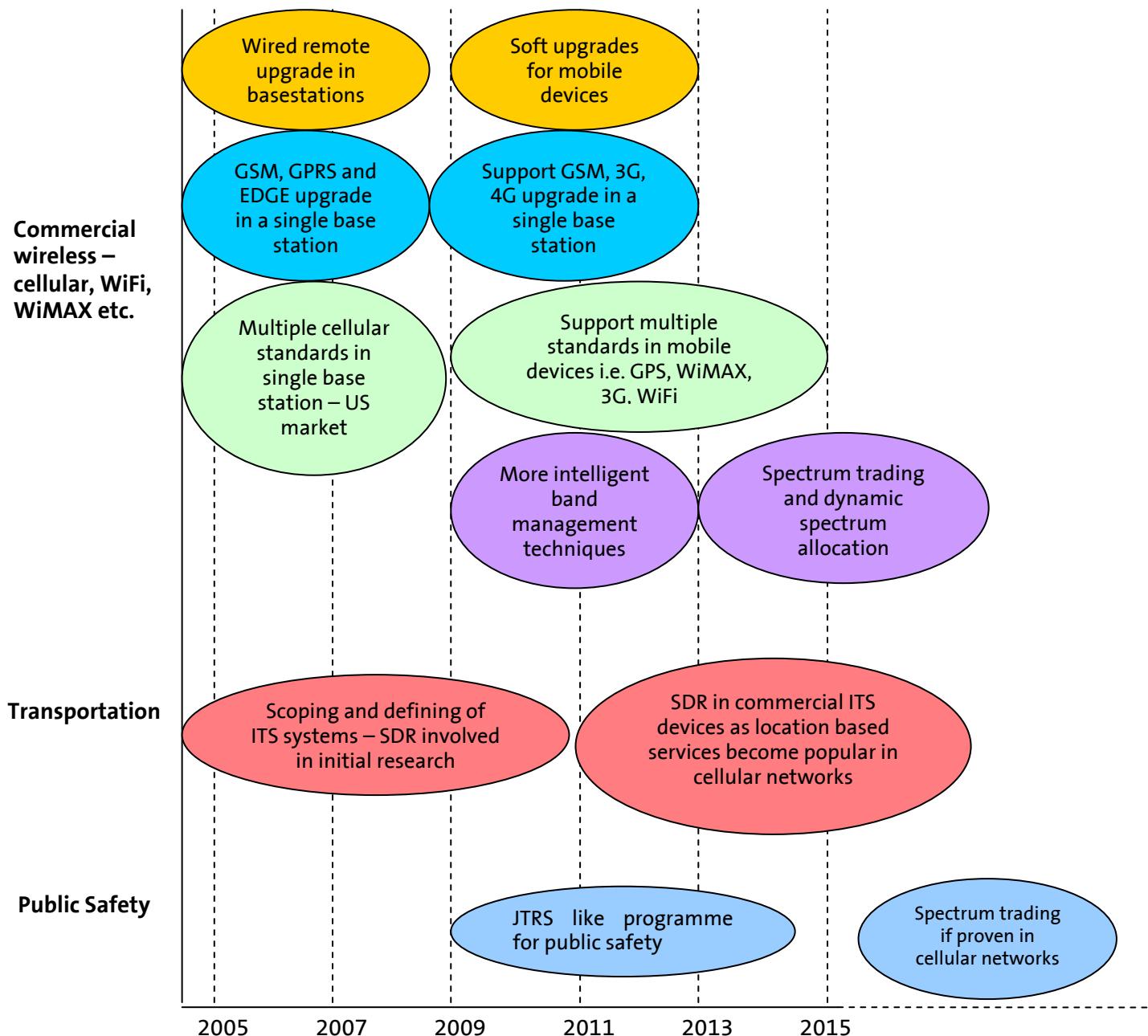


Figure 12-10: Forecast of future applications of SDR

12.6 Conclusions

This chapter has shown that, outside the defence sector, the commercial markets that SDR could benefit are:

- Cellular Networks
- Wireless Network Services i.e. WiFi, WiMAX and Bluetooth

UNCLASSIFIED

- Transportation
- Public Safety.

Of these, the cellular network industry is the market that is being targeted the most with SDR products already being deployed. This market is likely to continue to grow, especially as other wireless network services, such as WiFi and WiMAX, are adopted to complement existing cellular products. This will appeal to a wider variety of users. The ability of SDR to provide a smooth migration path and support multiple standards makes it suitable for this new mixture of cellular services.

SDR products are already on offer in the cellular network industry from companies like Vanu and AirNet. Base stations that can be remotely upgraded and support migration from GSM to GPRS to EDGE are available. The next step will be to provide base stations that can provide smooth migration to 3G and 4G services. While the technology to do this is available (experimental GSM combined with 3G systems show this) there are significant commercial barriers to overcome. Network operators' strong preference to deal with established manufacturers who can provide a complete network solution, as opposed to just the base station sub-system, could delay the widespread deployment of products from SDR start-ups. Interoperability with existing network infrastructures is also a difficulty.

The ability of SDR start-ups to continue to penetrate the cellular industry could be helped by standardisation initiatives, like OBSAI, and their ability to form partnerships with existing established manufacturers. It is clear that a general migration to reconfigurable platforms has begun, with the widespread usage of DSPs and FPGAs in 3G base stations. However, the pace of this migration could be driven by the success of SDR start-ups and pressure on established manufacturers to keep pace.

Due to the adoption of multiple cellular standards, the US will be the most attractive market for current SDR base stations. Once technical barriers are overcome to incorporate SDR into mobile devices, its adoption could become much more widespread. The ability of SDR to offer mobile users multiple services from one device could be the key factor that drives its widespread adoption. Due to power constraints, early mobile SDR devices may not be handsets but a size comparable to laptops or PDAs. These will migrate to handsets as technical barriers are overcome but this will take some years to achieve.

Lack of interoperability in the public safety sector, particularly in the US, means that it could really benefit from SDR. However, long procurement cycles in this sector mean that it will be a slow process to obtain widespread adoption of SDR. Investment in a public safety version of the JTRS programme to give incentives for developments in this area could really pay dividends for the public safety sector.

Finally, the growing interest in national and international compatibility in the transport sector could be a good opportunity for SDR. This is an emerging market that will take time to define itself. However, it does offer a market that SDR could dominate from the outset.

Overall, SDR fits with the long-term future direction of communications as the industry moves towards more complex and flexible systems incorporating more services. SDR is a key enabling technology of more spectrally aware devices, such as Cognitive Radio, which in turn could help reach targets for greater spectral efficiency and spectrum trading in the long-term.

13 Spectrum Efficiency Gains of Software Defined Radio and Cognitive Radio



By Tim James, Multiple Access Communications Ltd.

13.1 Introduction

Attention is placed on the spectrum efficiency gains of software defined radio (SDR) and cognitive radio (CR) in this chapter, and consideration is given to how CR-enabled SDRs might help improve the efficiency with which the RF spectrum in the UK is utilised. We begin by noting that some of today's radio technologies already exhibit limited CR characteristics. A few of the more prominent examples are discussed in Section 13.2. Three basic classes of CR are considered in the remainder of the report. Perhaps the simplest form of CR might consist of a multi-mode terminal that intelligently switches between modes according to service availability and the requirements of the user. This form of CR is discussed in Section 13.3. In Section 13.4 we consider CRs that use dynamic spectrum allocation (DSA) to take advantage of unused RF spectrum and in Section 13.5 we consider the idea of dynamically adjusting other physical layer parameters such as transmit power, modulation method and/or channel coding to maximise the efficiency with which spectrum is used. Our main findings are summarised in Section 13.6.

13.2 Existing Systems Exhibiting Limited Cognitive Capabilities

In this section we are primarily interested in the spectral benefits of CR-enabled SDR. However, as stated previously, SDR is not necessarily a prerequisite of CR. As proof of this, some existing radio systems already exhibit limited cognitive capabilities. Two of these, namely, digital enhanced cordless telecommunications (DECT) and universal mobile telecommunications system (UMTS), are discussed briefly in the following.

13.2.1 Digital Enhanced Cordless Telecommunications

DECT [366] is most commonly associated with its application in domestic cordless telephones. However, DECT is designed to support numerous applications and services and is part of the IMT 2000 third-generation (3G) family of cellular technologies.

DECT nominally operates in the 1880 to 1900 MHz band, employing time-division multiple access (TDMA) and time-division duplex (TDD) across ten RF channels [367]. DECT devices communicate using a Gaussian frequency shift keying (GFSK) modulation scheme. The channel bit rate is 1152 kbps and the transmit power is fixed at 250 mW (24 dBm).

UNCLASSIFIED

The cognitive element of the DECT radio system lies in its dynamic channel selection and allocation mechanism. The time domain is subdivided into 10 ms ‘timeframes’, each consisting of 24 timeslots. Nominally, the first 12 timeslots in each timeframe are used for downlink transmissions and the remaining timeslots are used for uplink transmissions. Thus, a DECT system can theoretically support up to 120 full duplex channels. DECT equipment periodically scans its local radio environment, measuring the received signal strength on all channels. The channels are then sorted based on the measured signal strength data; channels reporting ‘high’ signal strength are considered to be occupied; channels with low reported signal strength are considered to be idle. When a new channel is to be set up, the transmitting equipment will select what it considers to be the ‘best’ channel from its prioritised channel list.

The dynamic channel selection and allocation mechanism is not only used when a channel is first set up. If, during the lifetime of a radio link, the active channel becomes subject to interference, a duplicate channel can be set up. The base station will then monitor both channels before discarding the link with the lowest quality.

DECT is a good example of how a CR might use DSA to allow multiple, independent communications devices to coexist within shared spectrum. We note that, in spite of its cognitive tendencies, DECT is not a particularly new standard; the first edition of the DECT standard was released in 1992. Furthermore, we note that DECT does not require the use of SDR in its implementation.

Whilst the dynamic channel selection and allocation used in DECT allows it to coexist with other nearby DECT systems, it is perhaps important to note that an altogether more advanced system might be required if multiple systems were to share common spectrum. Multiple DECT systems can coexist because all systems share a common RF channel definition and a common timeslot structure. Because it is a multi-carrier system, DECT might tolerate interferers on individual carriers. However, on a single carrier, a DECT system cannot adapt its timing to coexist with other radio systems employing different timeslot structures. In the ultimate vision of CR, it might be desirable for multiple systems to be able to coexist.

13.2.2 Universal Mobile Telecommunications System (UMTS)

The UMTS terrestrial radio access (UTRA) radio systems, specifically, UTRA frequency division duplex (FDD) and UTRA TDD, are emerging as the dominant 3G cellular technologies in Europe. Operating in both paired and unpaired spectral bands around 2 GHz, UMTS uses code-division multiple access (CDMA) in 5 MHz channels.

Two key features of UMTS make it interesting when considering radios with cognitive capabilities. First is its use of transmit power control (TPC) to maintain a constant energy-per-bit to interference power spectral density, E_b/I_o , and second its use of variable spreading factors to allow user numbers to be traded off against data bandwidth.

UNCLASSIFIED

Unlike cellular systems such as GSM, a UMTS network can be deployed with only a single RF carrier. Thus, whereas in a properly planned GSM network the capacity of a cell is a function of the number of RF carriers assigned to it, capacity in a CDMA network is theoretically limited by the co-channel interference received from other users. Because of this, the transmit power of its users is critical; if the transmit power is too low, link quality will deteriorate to unacceptable levels; if the transmit power is too high, overall network capacity will be reduced because of unnecessary interference levels. So that transmit levels are maintained at optimum levels as users move about the network, UMTS implements a closed-loop TPC mechanism; by sending TPC commands back to the transmitter, the receiver can instruct the transmitter to adjust its output power in order to maintain an 'adequate' link quality. Thus, users located towards the edge of a cell will typically be instructed to transmit at a higher power than those closer to the base station.

Perhaps the most important point here is the concept that the radio uses a target E_b/I_o to control its transmit power. This is a good example of how a CR might maximise spectral efficiency. A more 'selfish' radio system might simply strive for optimum link quality by transmitting on full power. While this might offer incremental improvements in link quality for the entities involved, it could have a significant impact on the ability of other users to share and reuse the same spectrum.

From its inception, UMTS was designed to support a range of services, from speech to high-speed data transfer. UMTS uses a basic modulation chip rate of 4.096 Mcps. Channel data are 'spread' to the chip rate by multiplying the channel data by a 'channelisation' code. The spread channel data are then encoded through the application of a 'scrambling' code, with each base station and mobile station adopting a separate code. The channelisation (or orthogonal variable spreading factor (OVSF)) codes are organised such that the number of channels can be traded off against the data bandwidth of the individual channels. This ability to support multiple data rates with different length channelisation codes means that multiple users operating very different services can coexist. Note that, on the downlink, the organisation of the channelisation codes means that user numbers can be traded for fewer, higher bandwidth connections.

The use of variable length channelisation codes in UMTS gives an example of a radio system effectively modifying its waveform characteristics to adapt to the requirements of the chosen service. If CRs are to maximise the efficiency with which spectrum is utilised, such behaviour might be essential, especially as the requirement to support a wide range of data rates and service types becomes increasingly important. Note that there are many ways in which a waveform might be adjusted dynamically to optimally meet the requirements of the user in a variable RF environment. Two significant characteristics that might be altered are the modulation scheme and the forward error correction (FEC) levels.

UNCLASSIFIED

From its inception, UMTS was designed to support a range of services, from speech to high-speed data transfer. UMTS uses a basic modulation chip rate of 4.096 Mcps. Channel data are ‘spread’ to the chip rate by multiplying the channel data by a ‘channelisation’ code. The spread channel data are then encoded through the application of a ‘scrambling’ code, with each base station and mobile station adopting a separate code. The channelisation (or orthogonal variable spreading factor (OVSF)) codes are organised such that the number of channels can be traded off against the data bandwidth of the individual channels. This ability to support multiple data rates with different length channelisation codes means that multiple users operating very different services can coexist. Note that, on the downlink, the organisation of the channelisation codes means that user numbers can be traded for fewer, higher bandwidth connections.

The use of variable length channelisation codes in UMTS gives an example of a radio system effectively modifying its waveform characteristics to adapt to the requirements of the chosen service. If CRs are to maximise the efficiency with which spectrum is utilised, such behaviour might be essential, especially as the requirement to support a wide range of data rates and service types becomes increasingly important. Note that there are many ways in which a waveform might be adjusted dynamically to optimally meet the requirements of the user in a variable RF environment. Two significant characteristics that might be altered are the modulation scheme and the forward error correction (FEC) levels.

13.3 Multi-Mode Terminals

Perhaps one of the simplest forms of CR will consist of a multi-mode terminal with an intelligent control entity that can automatically switch modes depending on the location of the user, the user’s requirements and service availability. Thus, these so-called ‘always best connected’ (ABC) devices try to configure themselves to provide the user with an optimal compromise between connection performance, reliability and cost. By switching between the supported modes in a seamless manner, ABC devices have the potential to allow users to roam virtually anywhere without needing to be concerned with the availability of individual services.

SDRs are well suited to this kind of application because the same hardware might be reused to target the supported waveforms. Thus, an SDR-implemented terminal might be more compact and cheaper than a non-SDR terminal requiring physically separate signal paths for each standard supported. Furthermore, a programmable SDR might allow additional waveform ‘plug-ins’ to be downloaded to support new and updated radio standards as they become more widely adopted; a radio with a more traditional architecture would be unable to track evolving radio standards as easily, if at all.

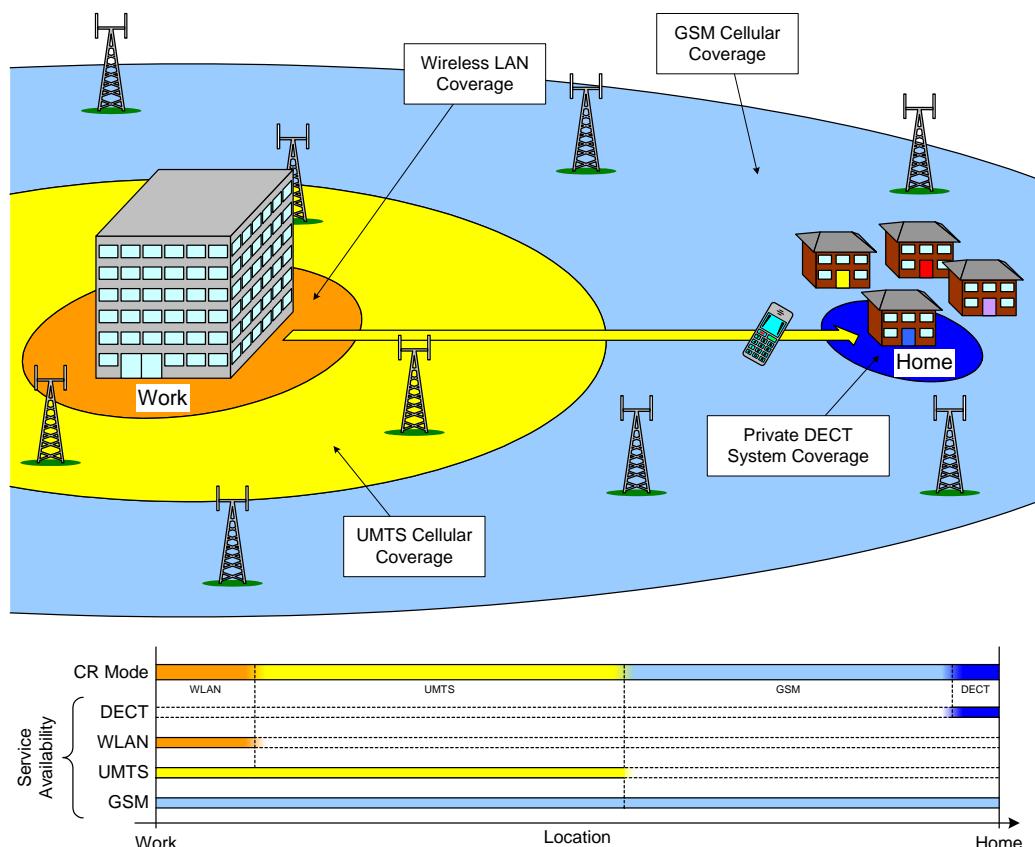
As an example, consider a terminal supporting DECT, GSM/EDGE, UMTS and a wireless local-area network (WLAN) technology such as 802.11a or 802.11b. The cognitive element of the terminal would monitor the requirements of the user, i.e., service type, and the availability of the supported services. Through the application of a simple ‘policy set’, the terminal would then automatically switch between the supported waveforms to provide the user with the most appropriate terminal configuration. For example, a simple policy might be to prioritise the waveforms as listed in Table 13-1.

UNCLASSIFIED

Priority	Waveform	Conditions
1	DECT	Only connect to recognised systems, e.g., home.
2	WLAN	Only connect to recognised networks, e.g., work.
3	UMTS	
4	GSM/EDGE	

Table 13-1: Simple waveform usage policy for a multi-mode CR

Consider now a scenario in which the user travels from his/her workplace to his/her home, as shown in Figure 13-1. When within the workplace, the terminal will automatically connect to the office WLAN; this connection might be used for both data and voice services using voice-over-internet protocol (IP) (VoIP), for example. On leaving the workplace the terminal will drop out of range of the WLAN and automatically reconfigure itself as a UMTS terminal, again offering both voice and data services. As the user moves away from the urban centre they might leave the coverage of the UMTS network. Without UMTS coverage the terminal will automatically fall back to operation as a GSM/EDGE terminal, offering voice and limited data services. Finally, on reaching his/her home, the terminal will detect the user's home DECT system and will reconfigure itself so that voice and data traffic can be routed via the user's landline connection.

*Figure 13-1: Operation of a multi-mode CR terminal*

UNCLASSIFIED

So what are the benefits of such a system? From the point of view of the user, such a device might save the user money by automatically taking advantage of more cost-effective connections when within range of recognised WLAN or DECT systems. Furthermore, support of GSM and UMTS cellular standards might provide the user with a higher probability of coverage when on the move. Whilst such connectivity might be achievable today, it invariably requires the carrying of multiple radios or the manual intervention of the user to switch between ‘modes’.

It is acknowledged that the large telcos are unlikely to be receptive to schemes that will draw custom away from their networks. However, cellular operators in the UK already offer WLAN services. Vodafone, for example, currently offer business users access to over 1500 WLAN hotspots across the UK (in association with BT Openzone) [368]. Their locations include airports, hotels, conference centres and motorway service stations. T Mobile are also actively deploying WLAN hotspots across the UK [369]. This clearly demonstrates that cellular operators have acknowledged the importance of WLAN systems and are prepared to invest heavily in the technology. If CR-enabled SDR platforms were able to facilitate seamless and transparent transitions between cellular and other technologies, the cellular operators would gain a novel method of relieving congestion in their networks.

Whilst it crucial that users can clearly see a significant benefit if such technology is to be accepted, in this study we are primarily interested in improving spectral efficiency. How might such a cognitive system help improve the efficiency with which spectrum is used? At first glance it might be difficult to see how such a system would help improve spectral efficiency because the supported waveforms do not differ from those implemented by contemporary, dedicated terminals. Where this kind of system has the potential to improve spectral efficiency lies in its pre-programmed desire to camp on short range systems operating in unlicensed spectrum. Thus, by off-loading users from wide-area cellular systems to local-area hotspots, the load on the cellular systems might be reduced greatly. This, in turn, will relieve congestion in the valuable, licensed cellular frequency bands.

In principle, therefore, we can say that multi-mode CR terminals may help improve spectrum utilisation. However, in order to place any qualitative figure on the amount by which efficiency might be improved would require the implementation of a complex simulation model. Furthermore, even with such a simulation model, any results are likely to be heavily influenced by any assumptions taken regarding system availability, traffic densities and traffic movement.

Finally, we note that services not too dissimilar to those described above are about to become a commercial reality. BT Group’s BT Fusion service [370][371][372] aims to combine the mobility of a cellular handset with the low cost of landline voice calls when at home. Under this service, handsets will route voice traffic via a broadband connection when within range of a so-called ‘BT hub’ installed in the user’s home using Bluetooth wireless technology. On leaving the coverage of the BT hub, the handsets will automatically handover to the local (Vodafone) cellular network. If this service proves to be successful, it may only be a matter of time before SDR technology is used to integrate additional waveforms within mobile terminals in the manner described above.

13.4 Dynamic Spectrum Allocation

In the previous section we considered a simple form of CR based around existing wireless standards. This represents a very limited, simplistic application of CR. By its incorporation into new wireless systems, CR has the potential to enable significantly greater performance gains. An application of CR that is widely discussed is its application to enabling dynamic spectrum allocation. DSA attempts to maximise spectral efficiency by allowing devices to use unused spectrum on a short-to-medium-term basis. DSA might be implemented in both licensed and unlicensed spectrum. Both applications are discussed in more detail in the subsequent sections. SDR is often discussed in the context of CR-enabled DSA because SDRs are often associated with radios that are extremely frequency agile at RF as well as highly flexible at base band.

13.4.1 How Can DSA Help Improve Spectrum Utilisation?

In general, under the current regulatory environment, access rights to licensed spectrum are closely controlled, with only primary licence holders allowed to operate. Furthermore, the systems and services that may be operated in each band are strictly regulated. This often leads to the inefficient utilisation of spectrum for three reasons:

- Firstly, a user allocated a block of spectrum might only use his/her spectrum in selected geographical regions. Consequently, this spectrum might remain unused in large regions of the country.
- Secondly, an operator might seek to reserve access rights for more spectrum than they actually need, thereby protecting his/her ability to ‘grow’ into the unused spectrum at a later date.
- Finally, depending on the service type, demand for spectrum can fluctuate on a short-term basis. Under the current method of spectrum allocation, operators must dimension their spectral needs based on peak network load. Consequently, large blocks of potentially valuable RF spectrum might remain completely unused for a significant part of each day.

As a simple example, activity across a 20 MHz band centred on 902.5 MHz was monitored over a 24 hour period. This band is allocated to the GSM 900 uplink. The receiver for these measurements was located at Multiple Access Communications Ltd’s (MAC Ltd’s) premises on a science park on the outskirts of Southampton. The results of this survey are shown in Figure 13-2. The time of day is shown on the vertical axis; the survey started at 9 am on the first day and ran for 24 hours. Frequency is shown on the horizontal axis and peak received signal strength indication (RSSI) is represented using colour. As an aside, evidence of the frequency hopping employed by certain GSM operators can be seen in Figure 13-2. For example, simultaneous activity can be seen on 903.8 MHz (Channel 69) and 904.4 MHz (Channel 72). There is also simultaneous activity on 901.8 MHz (Channel 59) and 912.2 MHz (Channel 111).

UNCLASSIFIED

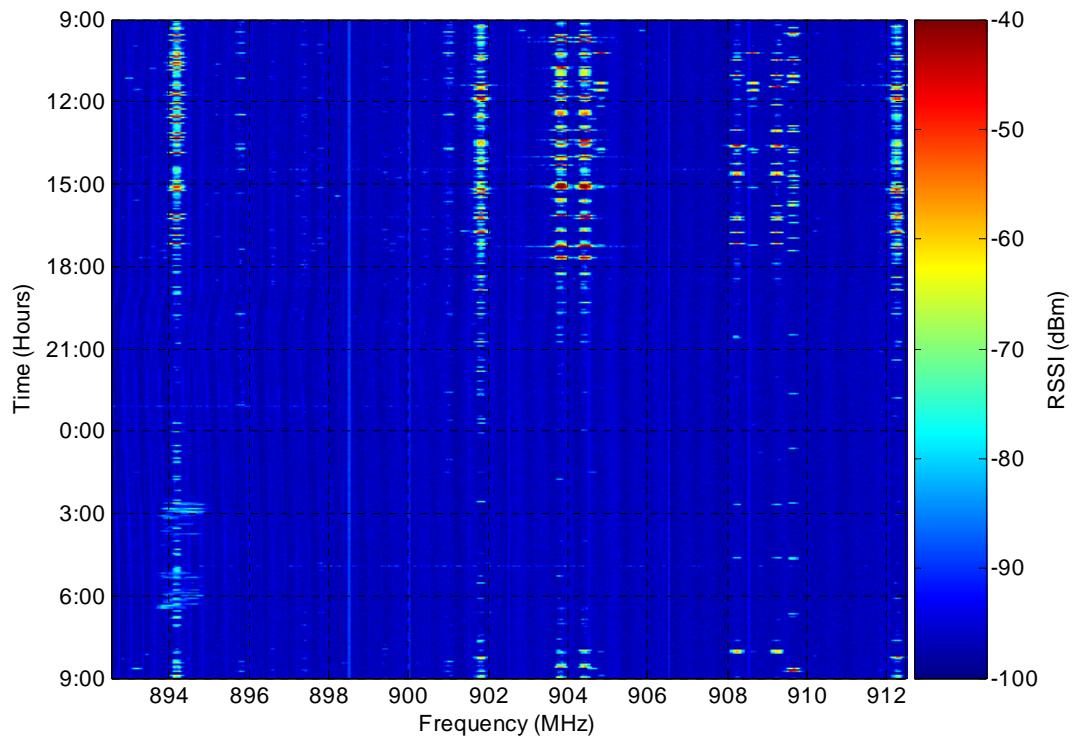


Figure 13-2: RSSI measured over a 24-hour period for a 20 MHz band of the GSM 900 MHz uplink band

A cursory glance at Figure 13-2 leads to a number of interesting observations:

- The majority of RF activity occurs between 8 am and 6 pm. Between 9 pm and 6 am there is very little activity. (This is perhaps not too unexpected given the location of the receiver, i.e., on a science park.)
- Much of the spectrum is ‘empty’. It is important to note that the spectrum picture at the base station antenna will be quite different from that shown above; the channel occupancy is likely to be significantly higher when viewed from the base station. However, despite these limitations, our measurements still suggest that there are large portions of the radio spectrum that remain unused.
- Even during the day, ‘occupied’ channels are unused much of the time. This is clearly evident in Figure 13-3, which shows the same RSSI data for the period 12 pm to 2 pm only.

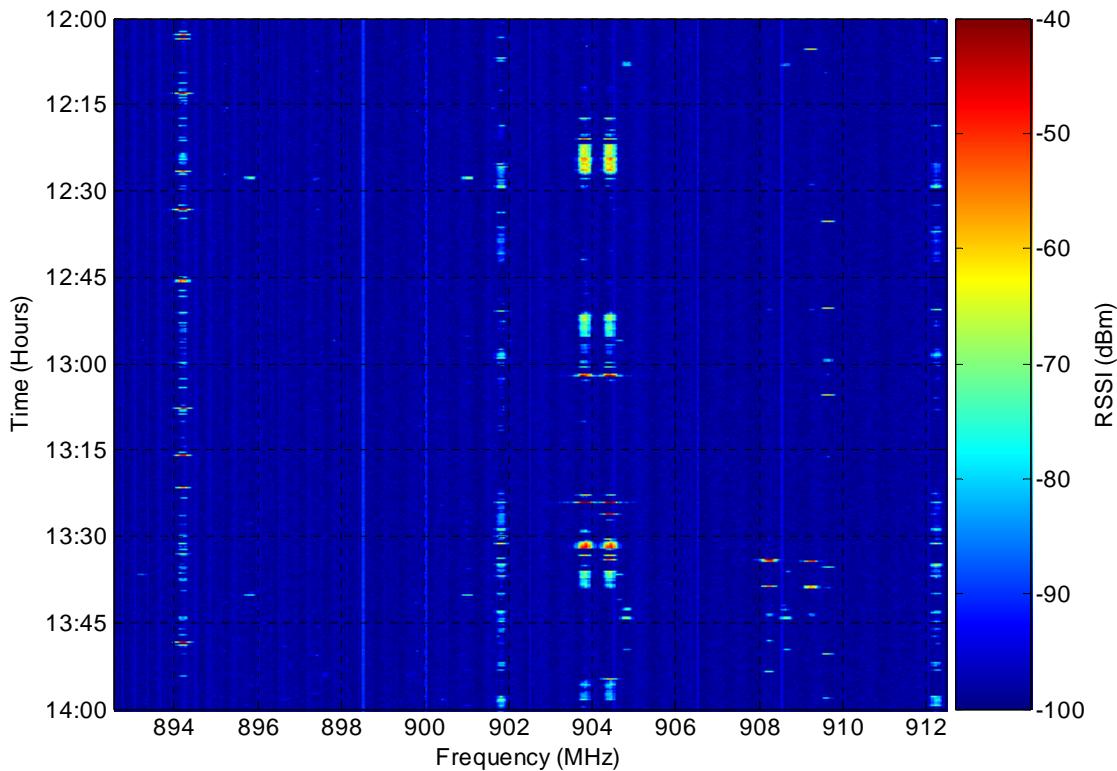


Figure 13-3: RSSI data from Figure 13-3 for the period 12 pm to 2 pm

From the example shown above, it is clear that a significant proportion of the available RF spectrum exhibits a low utilisation. Note that this scenario is not specific to the GSM 900 MHz uplink; large bands of frequencies are often free of significant radiated power right across the RF spectrum.

We have demonstrated using example RSSI data that in practice much of the available RF spectrum is not heavily utilised. Is there a realistic, theoretical limit to how efficiently spectrum might be used? And if so, might short-term DSA be used to improve this efficiency?

We begin by using the Erlang B formula [373] to analyse the relationship between blocking probability and mean channel utilisation. For reference, the Erlang B formula may be expressed as:

$$B = \frac{A^N / N!}{\sum_{k=0}^N \frac{A^k}{k!}} \quad \text{Equation 13-1}$$

UNCLASSIFIED

where A represents the offered traffic in Erlangs, B represents the probability that a call attempt is blocked and N represents the number of physical channels. If the carried traffic is given by $A \times (1 - B)$ and the theoretical capacity of the system is N Erlangs (i.e., all N channels are used 100% of the time), the mean channel utilisation is given by:

$$B = \frac{A^N / N!}{\sum_{k=0}^N \frac{A^k}{k!}} \quad \text{Equation 13-2}$$

Figure 13-4 shows the mean channel utilisation as a function of blocking probability for systems with between five and 200 physical channels.

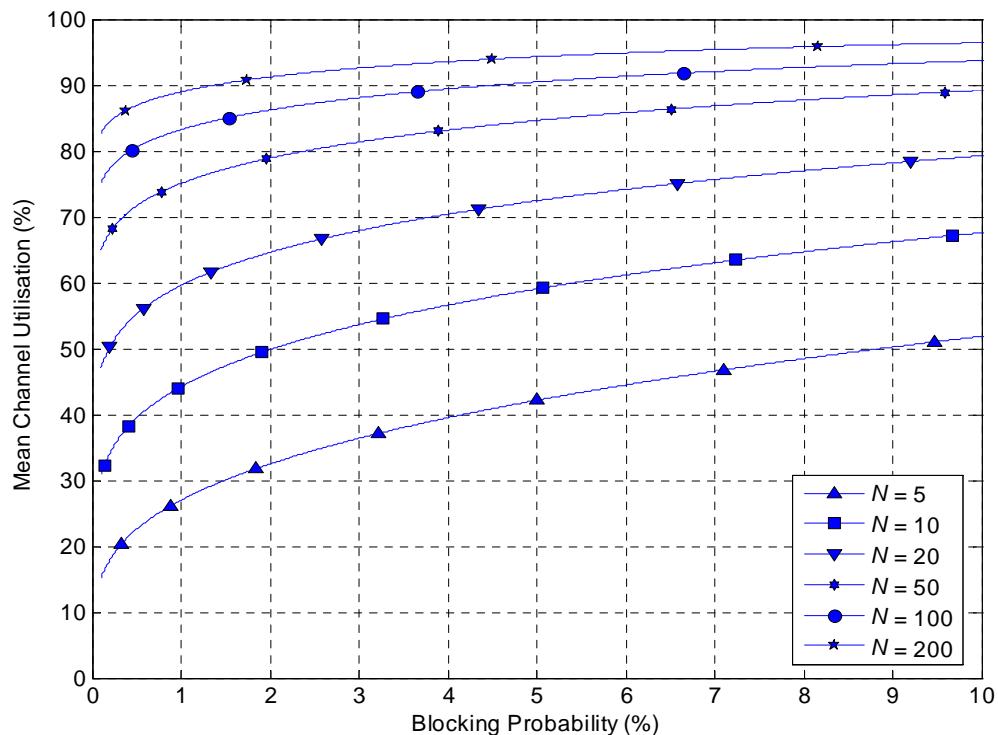


Figure 13-4: Relationship between blocking probability and mean channel utilisation for systems with between five and 200 physical channels

It can be seen that a significant proportion of the physical channels are, on average, unused. For example, in a system with 50 channels designed for a blocking probability of 2%, mean channel utilisation is approximately 79%. In other words, 21% of the physical channels are unused on average. Moreover, assuming that the network is dimensioned for peak loading, this percentage will increase as the offered traffic, i.e., load, is decreased.

UNCLASSIFIED

From Figure 13-4 we can see that mean channel utilisation for a given blocking probability appears to improve as N increases. This relationship is shown more clearly in Figure 13-5 in which we consider the relationship between mean channel utilisation and the number of physical channels, N , assuming various blocking probabilities. It can be clearly seen that, for a given blocking probability, mean channel utilisation tends towards 100% as N increases.

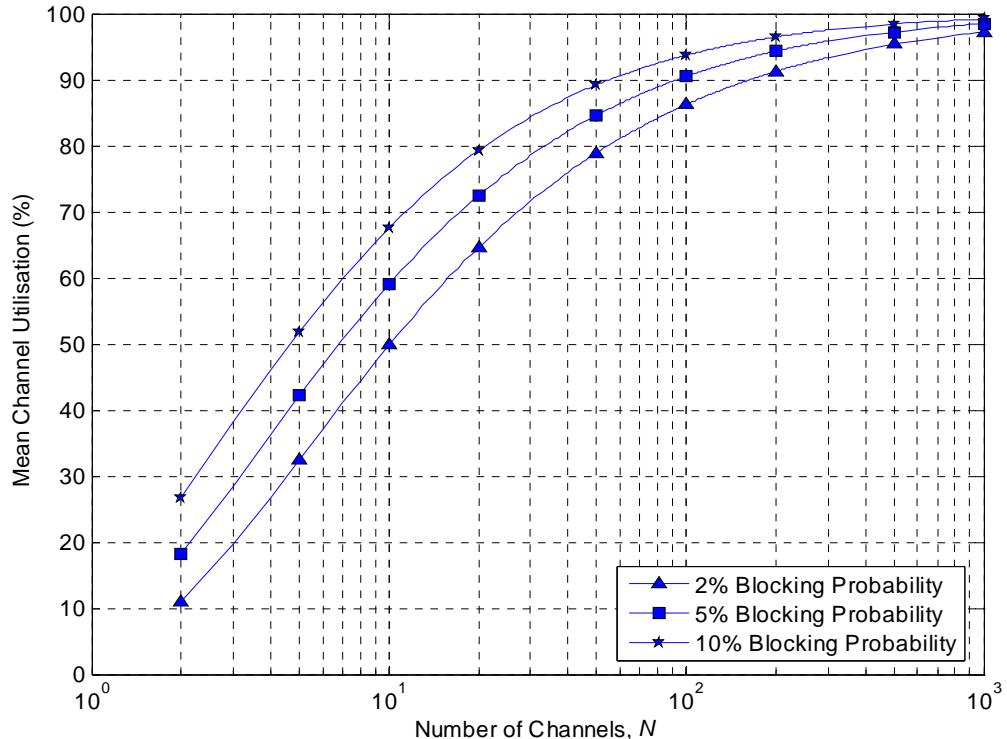


Figure 13-5: Relationship between the number of physical channels, N , and mean channel utilisation for a range of blocking probabilities

The implication of this is that greater overall spectral efficiency might be achieved by allowing multiple users to share a common ‘pot’ of spectrum (e.g., using DSA techniques) rather than breaking the spectrum up and allocating each user a fixed number of channels. For example, with a 2% blocking probability, a system with 100 channels can carry 33% extra traffic than five independent systems each with 20 channels (note that this analysis supports the ‘bank cashier’ analogy presented in Chapter 9).

UNCLASSIFIED

13.4.2 Brokered DSA in Licensed Spectrum

The proposed relaxation on spectrum trading and spectrum liberalisation brings with it the possibility that primary licence holders might allow third parties to use parts of 'their' spectrum on a temporary basis. It is reasonable to assume that primary licence holders are unlikely to want third parties accessing spectrum for which they have paid licence fees for free. Therefore, a method of billing would be required. In return the lessee is likely to expect certain guarantees regarding nominal interference levels, for example.

The first form of DSA to be considered therefore is that of 'brokered' DSA in licensed spectrum. Here, an on-line broker acts as a central agent to control the temporary allocation of spectrum in licensed bands. The broker might act on behalf of multiple licence holders and would be responsible to maintaining a database of available spectrum in both space and frequency domains based on data provided by the licence holders regarding spare spectral resources. The broker might also handle spectrum billing on behalf of its clients.

This form of DSA should have advantages for both the primary licence holder and potential lessees; the licence holder has the potential to gain extra revenue through the leasing of unused spectrum whilst spectrum users with limited needs are presented with a more cost-effective path to spectrum access without having to pay significant sums for exclusive access rights. There is also a business opportunity here for the role of the spectrum broker, who might expect to charge management fees and/or receive a proportion of the billing revenue.

An example of how such a system might operate is shown in Figure 13-6. We start with the user (shown on the left) wanting to make a call. As with existing systems, the user makes a random access attempt to request a data channel. Note that in order for the base station to reliably receive incoming random access attempts, mobiles would need to transmit on a predetermined set of channels. If, on receiving the random access message, the base station determines that it does not have access to sufficient spectrum to support a new data channel, the base station accesses the on-line broker (shown on the right) to negotiate access to dynamically allocated spectrum. The base station would send location data to the broker along with the key characteristics of the intended waveform, e.g., bandwidth and maximum transmit power. The broker would then negotiate with the base station to find a suitable slice of RF spectrum. On completion of this, the broker would send confirmation to the base station. This might include conditions of use, e.g., maximum transmit power, geographic constraints and expiry time. Once temporary spectrum had been allocated, the base station can then direct the mobile to open a data channel in the newly acquired spectrum. Spectrum might be allocated on a timed basis, i.e., once a band of spectrum is allocated it is flagged as 'occupied' for a predetermined time. Alternatively, on termination of the data transfer, the base station might inform the broker that the spectrum has been released so that a) billing can stop and b) the spectrum can be reallocated. The latter might ultimately provide the highest spectral efficiency although the first might be simpler to implement and manage.

UNCLASSIFIED

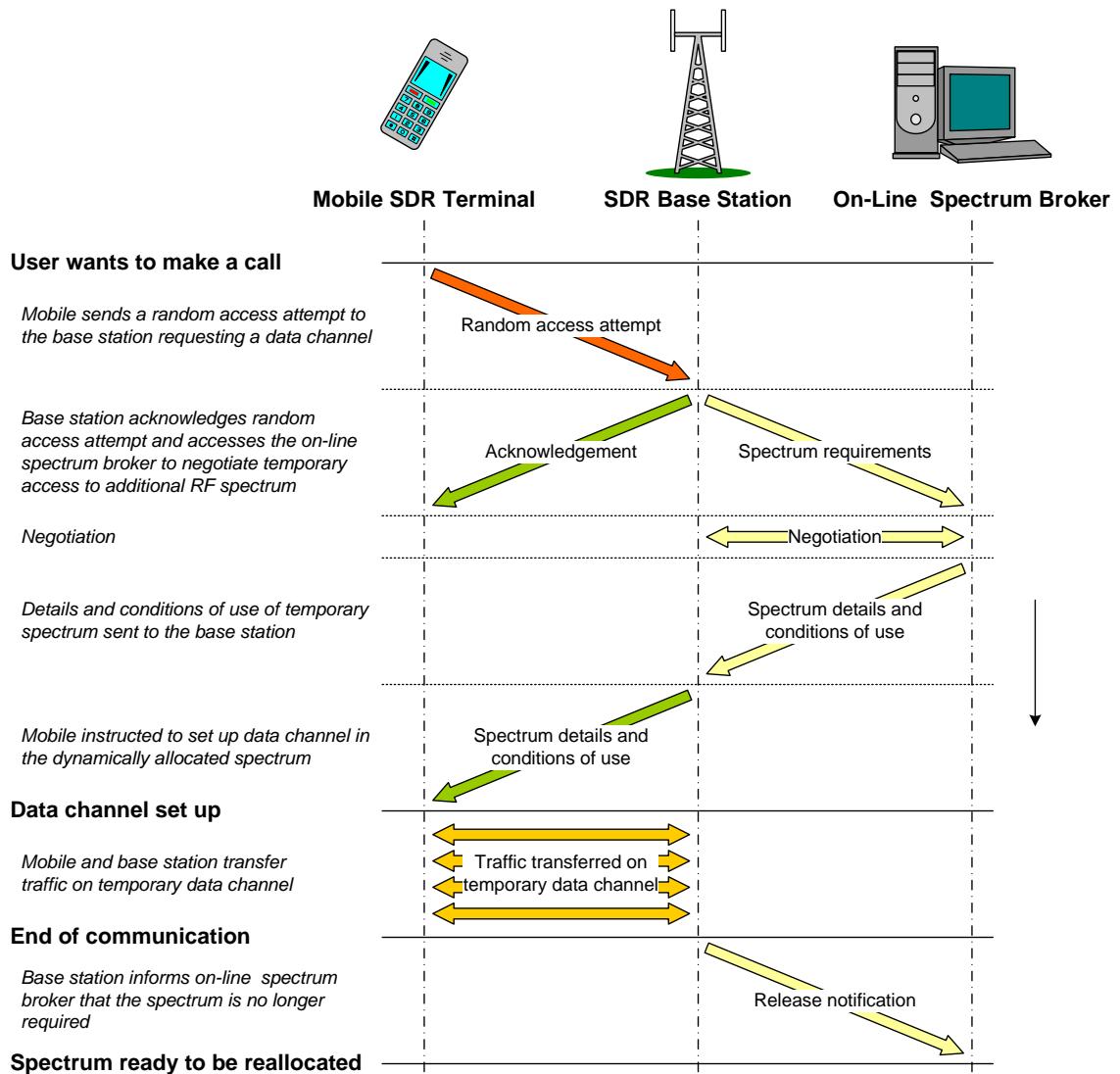


Figure 13-6: Example showing how temporary data channels might be set up under a brokered DSA scheme

In the example above we have shown the mobile initiating a call. Of course, the reverse is also possible with the base station paging the mobile to find its location before requesting spectrum for the necessary traffic channels.

To allow spectrum to be reused efficiently, the broker will have to process location and waveform data supplied by the spectrum users to maintain adequate user-to-user separation. For example, frequency bands might be reused more often, i.e., with lower user-to-user separation, for low-power users compared to that which can be achieved for high-power users; consequently, it might be reasonable to expect high-power users to be charged more per megahertz per minute than low-power users. Managing spectrum reuse would therefore require the application of a comprehensive policy set to ensure that spectrum is allocated reliably and fairly.

UNCLASSIFIED

Significant effort has been expended investigating dynamic spectrum allocation as part of the Information Society Technologies (IST) Dynamic Radio for IP-Services in Vehicular Environments (DRiVE) and OverDRiVE projects. In particular, these projects considered DSA as a means of improving spectral efficiency in a scenario in which two services, namely, a cellular UMTS service and a digital terrestrial TV (DVB T) service, share common spectral resources. Leaves et al [274] suggest that the period of peak demand for spectrum for these services differs. This is shown in Figure 13-7 [274], where the normalised demand for spectrum for a UMTS service is shown together with that for a DVB T service over a 24 hour period.

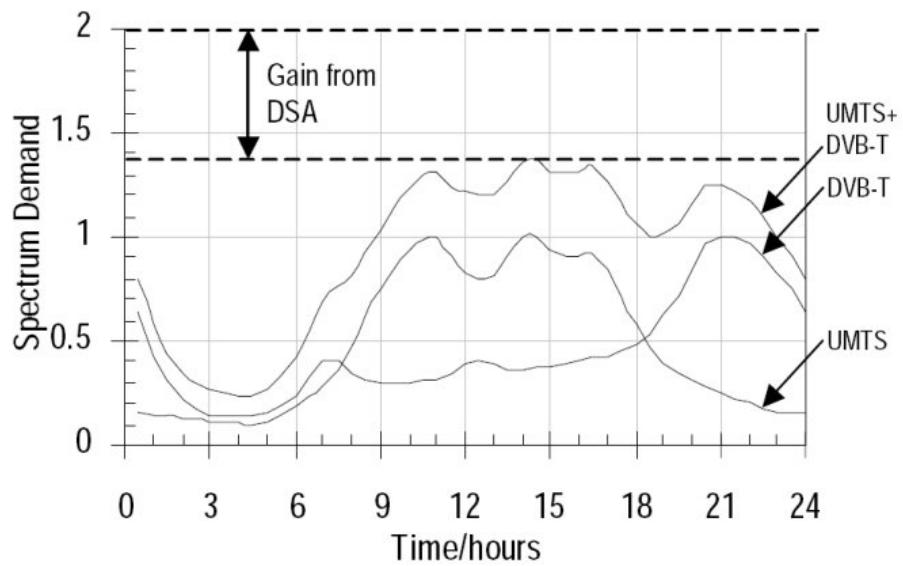


Figure 13-7: Example of DSA in unused DVB-T spectrum [274]

The peak *combined* spectrum demand for the two services is shown to be approximately 30% less than that required to support both systems in separate, statically assigned bands. Leaves et al conclude that if both services could share a common ‘pot’ of spectral resources rather than each being allocated a fixed bandwidth dimensioned according to worst case demand, a clear improvement in spectral efficiency could be achieved.

Although these data illustrate the potential spectral efficiency gains to be had through the adoption of DSA, we note that its performance is dependent on the lack of correlation between the spectrum demand profiles for the relevant services. In further work by Leaves et al [374], the relationship between the correlation between the service spectrum demand profiles and the DSA gain was considered. Figure 13-8 [374] shows how DSA gain decreases as the correlation between the relevant services increases. A negative traffic correlation value represents a low correlation between the spectrum demand profiles for each service. A traffic correlation of one represents identical spectrum demand profiles. The datum shown with a traffic correlation of approximately -0.2 is generated using the spectrum demand profiles given in Figure 13-7.

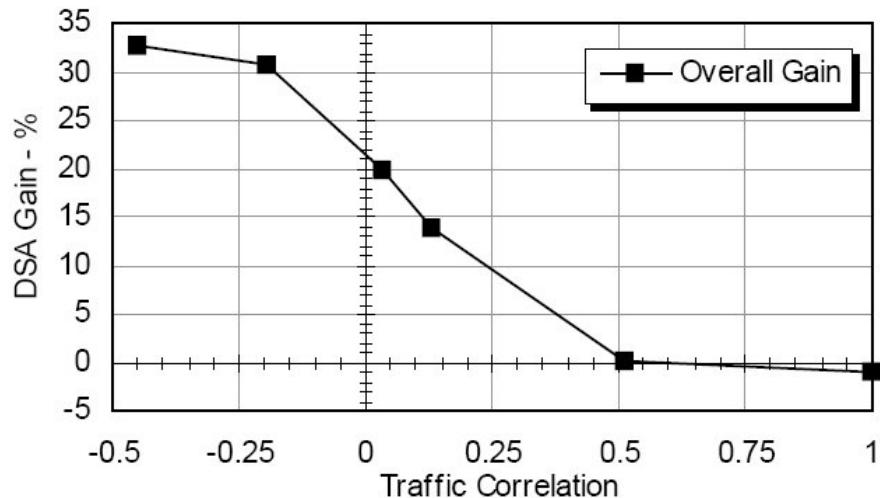


Figure 13-8: Simulated DSA gains as a function of service traffic correlation [374]

It is perhaps debatable whether the concept of brokered DSA is an application of CR; the mobile simply acts as a slave to the base station, with much of the ‘intelligence’ located in the base station and, in particular, the on-line broker. However, considering the radio system as a whole, this approach uses knowledge of the radios’ locations and capabilities to negotiate spectrum access with an independent agent. At a system level, therefore, this system does indeed demonstrate cognitive behaviour.

Brokered DSA requires a central agent to handle spectrum management and billing. Therefore, brokered DSA might be well suited to use by client/server networks, i.e., cellular networks, where the mobile users are slave to the network infrastructure. Such an approach is unlikely to be suitable for use by peer-to-peer users where access to an on-line broker might not be possible.

13.4.3 Opportunistic DSA in Unlicensed Spectrum

In the previous scenario we have considered the possibility of DSA in medium-to-long-range (typically client/server) systems operating in licensed spectrum. A significant opportunity for CR-enabled DSA lies with short-range devices operating in unlicensed spectrum (transmit power constraints for transceivers operating in unlicensed spectrum generally limit its use to short-range applications). Here, devices will combine sensing technology with usage policies to find and utilise unused spectrum. Calculations completed by Ofcom suggest that 800 MHz of licence-exempt spectrum should be sufficient to support short-range services with a capacity of 100 Mbps per person [8]. CR-enabled devices might hold the key to unlocking the potential of this spectrum.

There already exist a number of systems that operate in unlicensed spectrum today. These include DECT (discussed briefly in Section 13.2.1), the 802.11 family of WLAN technologies and Bluetooth radio technology to name just three. Each of these systems adopts a different approach to multiple access, i.e., spectrum sharing.

UNCLASSIFIED

As discussed previously, DECT uses DSA to find unused TDMA channels. The primary method of multiple access employed by 802.11 systems is that of carrier sense multiple access with collision avoidance (CSMA/CA). Here, transceivers will ‘listen’ before transmitting. If another transmission is detected, the radio will delay for a random period before trying again. Finally, Bluetooth wireless technology takes an altogether less sophisticated approach. Bluetooth uses a frequency hopping spread spectrum (FHSS) approach with a short dwell time (625 µs) [375]. Here, radios hop in frequency using a pseudo-random sequence known by all members of a given ‘piconet’. Collision avoidance is not implemented. With Bluetooth wireless technology it is accepted that collisions between packets from different transmitters will occur occasionally. However, the resulting dropped packets can be tolerated because the interfering radios will normally jump to different frequencies before transmitting subsequent packets. (For data services, dropped packets can be retransmitted.)

Each of the systems mentioned above has been designed to coexist with other examples of the same system. However, the ability of the individual waveforms to coexist with one another is not clear. DECT, for example, relies on knowledge of the channel structure (in both frequency and time domains) to be able to scan and select unoccupied channels. Thus, DECT might have problems coexisting with non-DECT waveforms. Bluetooth, on the other hand, relies on frequency hopping to keep the probability of packet collision to a minimum and does not perform any ‘carrier sensing’. Thus, whilst Bluetooth might tolerate the occasional dropped packet, the same might not be true were Bluetooth to be operated with other short-range systems. Therefore, CRs developed to operate in unlicensed spectrum not only have to be tolerant of transmissions from ‘non-native’ radio systems, they will also have to actively avoid causing unacceptable interference to these same systems. In particular, CRs introduced into spectrum to be shared with existing legacy systems must be able to detect and work around these systems. In other words, access to unlicensed spectrum needs to be structured on the basis of fair play and not simply the survival of the fittest (i.e., who can shout the loudest). This is why the use of policies and the concept of radio etiquette are important.

13.4.4 Opportunistic DSA in Licensed Spectrum

As stated previously, it is reasonable to assume that primary spectrum licence holders will, in general, be unlikely to accept other users freely accessing spectrum for which they have paid licence fees. However, there may be instances in which limited, opportunistic access by relatively low-power, unlicensed CR devices might be tolerated. One such example is the use of unlicensed devices in unused TV spectrum [376]. Here, CRs might use spectrum monitoring techniques to identify and use ‘unused’ spectrum.

UNCLASSIFIED

Allowing CRs to be introduced into spectrum shared with receive-only equipment (e.g., TV and broadcast radio receivers) must be considered carefully. A CR using opportunistic DSA needs to ensure that its own transmissions will not cause unacceptable interference to other receivers operating nearby. Devices operating TDD waveforms effectively broadcast their presence through their transmissions. However, receive-only equipment and, perhaps to a lesser extent, devices operating FDD waveforms might not be detected so easily. Therefore it might not be sufficient to assume that if a CR is out of range of a distant signal that it is free to transmit because its transmissions might still cause interference to a receive-only device that is within range of the distant transmitter. In the case of a TV receiver this issue might be compounded due to the nominally high gain of a TV aerial, i.e., an acceptable TV signal received by a TV antenna mounted at roof height might be attenuated significantly when received through a 0 dBi CR antenna located closer to ground level.

Therefore, an obstacle to opportunistic DSA in licensed spectrum might be the reliable detection of priority transmissions, i.e., relying simply on power measurements might not be sufficient because it might be necessary to be able to detect transmissions received with negative SNRs. One approach to the detection of ‘buried’ signals is the use of so-called ‘cyclostationary’ detectors [377]. This form of feature detector uses long measurement times together with signal processing to achieve signal sensitivities below the noise for signals with published characteristics. It is reported that such techniques might allow signals more than 30 dB below the receiver noise floor to be detected [376].

13.5 Dynamic Power Control, Modulation and Channel Coding

DSA aims to improve spectral efficiency by ‘packing’ multiple users into shared spectrum. DSA requires flexibility in the RF front end. Another approach to improving spectral efficiency is to dynamically adjust the characteristics of the transmitted waveform to allow more efficient utilisation of the available spectrum. Two different approaches are considered in the following sections, namely, the use of dynamic power control and the use of dynamic modulation and/or channel coding.

13.5.1 Dynamic Power Control

We start by considering the application of dynamic power control. Dynamic, closed-loop power control or TPC is a well established concept. Indeed, TPC is a fundamental part of cellular CDMA systems such as UMTS (see Section 13.2.2). So, what is the point of TPC?

A radio receiver will typically require a nominal received signal-to-noise ratio (SNR) (e.g., stated in terms of the energy-per-bit-to-noise spectral density (E_b/N_0)) in order to satisfy a given bit error rate (BER) requirement. The BER requirement may be related to the current service type. The SNR requirement will take into account a number of factors. These will include the noise figure and noise bandwidth of the receiver, modulation method, forward error correction (FEC) scheme and the characteristics of the radio channel.

UNCLASSIFIED

Figure 13-9 shows theoretical BER curves as a function of SNR for both uncoded BPSK and uncoded QPSK systems for both Gaussian and Rayleigh radio channels. In order to achieve an uncoded BER of 10^{-2} (i.e., one error in every 100 bits) with a BPSK modulation scheme in a Gaussian channel, a SNR of 6.8 dB is required. Move to a Rayleigh fading channel and the mean received signal power must be increased by over 17 dB to achieve the same BER. In both cases, an additional 3 dB is needed to achieve the same BER with a QPSK modulation scheme. We note that this example is for uncoded systems. Adding FEC to the transmitted data bits will permit operation with a lower SNR. However, this will also reduce the effective number of user data bits per symbol because of the added overhead of the FEC data.

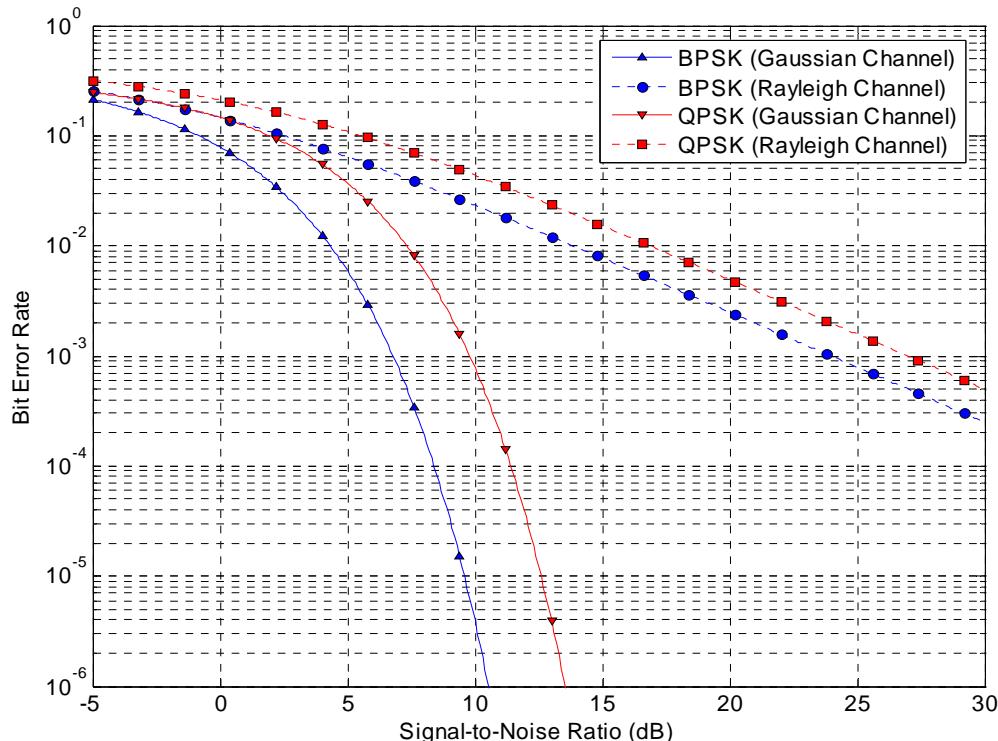


Figure 13-9: Theoretical BER curves for uncoded BPSK and QPSK modulation schemes in both Gaussian and Rayleigh fading channels [176][378]

From Figure 13-9 it can be seen that transmit power might be dynamically adjusted to maintain a ‘target’ BER (and hence received SNR). The target SNR is dependent on the current waveform. Furthermore, it can be seen that the target SNR might be adjusted according to the characteristics of the radio channel; for example, target SNR might be reduced significantly if the receiver can determine that the received signal is relatively free from the effects of fast fading (e.g., in the scenario in which the receiver has a line-of-sight link from the transmitter).

We have discussed how dynamic power control might operate. Other than offering potential savings in power consumption (and hence improved ‘talk time’), dynamic power control offers little benefit to the isolated user whilst potentially increasing system complexity significantly. What then are its benefits to other spectrum users?

UNCLASSIFIED

In order to maximise spectral efficiency in the spatial domain it is desirable to minimise the distance between independent users operating on the same frequency. The minimum separation between two users will typically be limited by co-channel interference. Thus, in order to reduce the user-to-user separation the carrier-to-interference-plus-noise ratio (CINR) must be reduced. Whilst one user might attempt to improve CINR by *increasing* transmit power, co-channel users might have to respond by increasing their own transmit power. Thus transmit power might snowball until the maximum transmit power is reached without any appreciable improvement to CINR overall. If, however, all users reduce their transmit power to an ‘acceptable’ minimum, mean interference power might be reduced overall.

Consider the simple example shown in Figure 13-10. In Figure 13-10 (top left) 80 ‘users’ are distributed on a rectangular grid within a 1 km radius of a ‘base station’. Here all users transmit with a fixed transmit power. The RSSI assuming a path loss coefficient of 3.5 is plotted for the surrounding area. In this example, the uplink signal received at the base station from close-in users is significantly greater than that received from users located at the ‘cell’ edge.

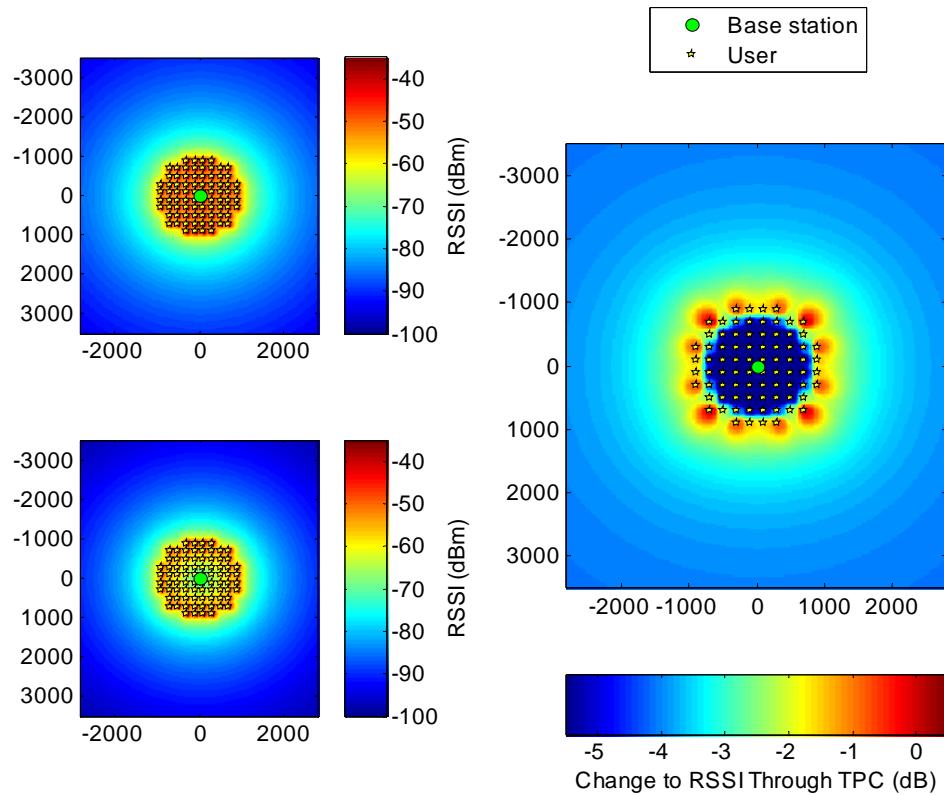


Figure 13-10: Example to illustrate spectral efficiency gain through dynamic uplink power control with 80 users distributed in a 1km radius of a ‘base station’; RSSI for all users transmitting ‘maximum’ power (top left), and all users adjusting their transmit power so that the power received at the base station is constant (bottom left), and the effect is an overall reduction in RSSI outside the ‘cell’ (right)

UNCLASSIFIED

Consider now the scenario in which all users adjust their transmit power so that the uplink signal power received at the base station is equal for all users (we assume that the users located at the cell boundary transmit the same power as before). The resulting RSSI for the surrounding area is shown in Figure 13-10 (bottom left). If we compare the uplink RSSI plots with and without TPC (shown in Figure 13-10 (right)), we find that on average the interference caused to co-channel users outside our cell is reduced by approximately 4.1 dB. Applying the same process to the downlink yields a 4.3 dB reduction in interference power.

If co-channel interference caused to others in an interference limited scenario is reduced by 4 dB, neighbouring systems might be able to operate with approximately 4 dB less path loss between systems. With a path loss coefficient of 3.5, a 4 dB reduction corresponds to a 23% reduction in range. An ability to reduce the separation between independent spectrum users by 23% implies that the number of users can potentially be increased by almost 70%, a significant improvement. We note that this is a simplistic example, possibly yielding an overly optimistic figure. We have assumed perfect power control, for example. However, this example clearly demonstrates the concept. As mentioned in our earlier discussions, further gains might be made by dynamically adjusting target SNR according to the characteristics of the radio channel (in our example we have effectively chosen a constant target SNR).

13.5.2 Dynamic Modulation and/or Channel Coding

In the previous section we considered using dynamic power control to help maximise spectral efficiency in terms of minimising the distance needed between independent radio systems sharing common spectrum. Now we move to consider the application of dynamic modulation and/or dynamic channel coding. Here the focus is rather on using improved SNR nearer the base station to maximise channel throughput.

Consider the simple example shown in Figure 13-11. The base station in Figure 13-11 (top) uses a binary modulation scheme, e.g., BPSK, to transmit and receive data from users operating within its coverage. For a given QoS, a maximum BER and hence minimum E_b/N_0 or SNR is required. Assuming a constant transmit power, we note that the effects of path loss mean that user's towards the centre of the cell operate with an unnecessarily high SNR (as shown in Figure 13-11 (bottom)) – this implies an inefficient use of the spectrum.

UNCLASSIFIED

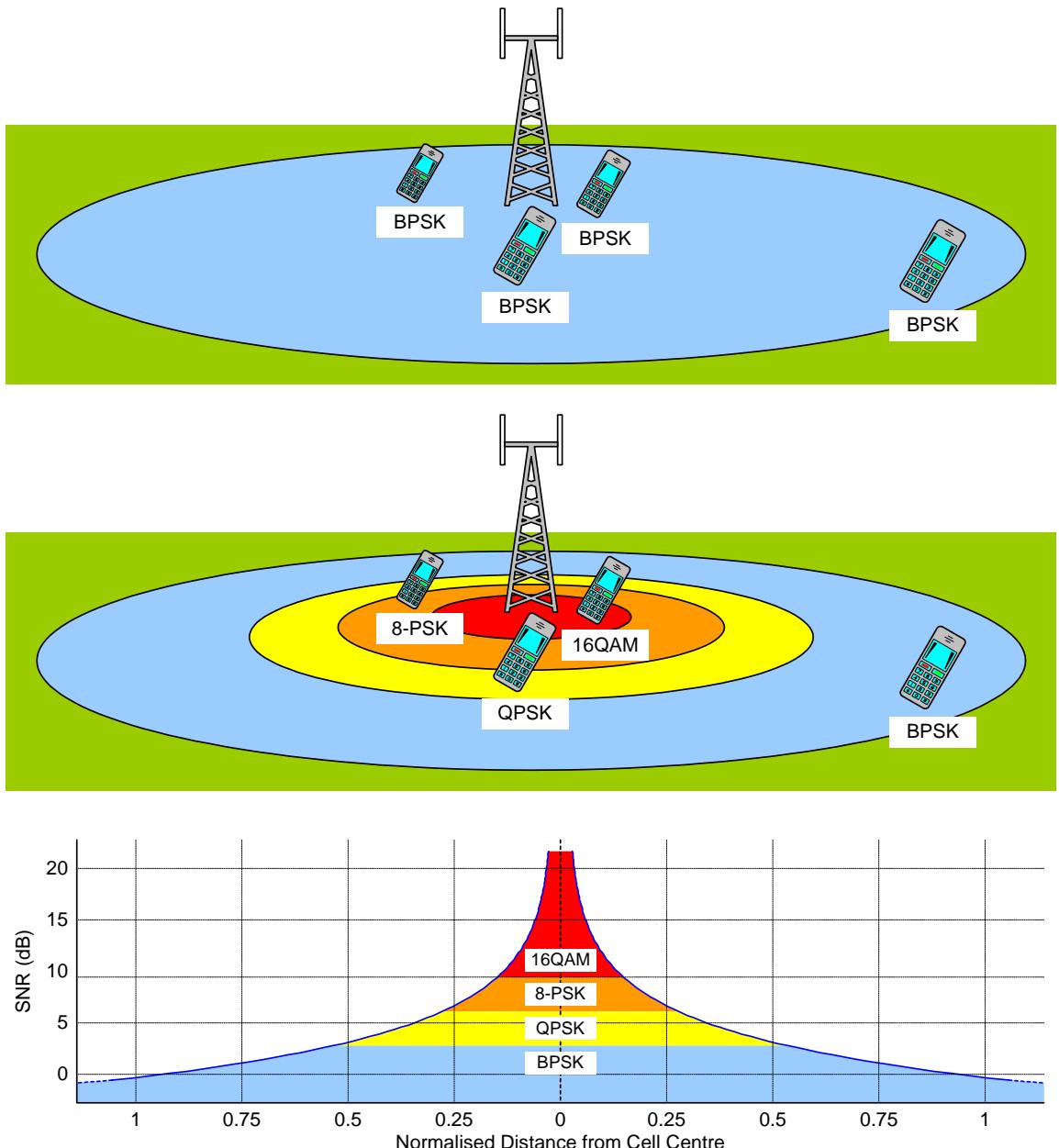


Figure 13-11: Using dynamic modulation to improve spectral efficiency of a cellular network; an isolated cell without dynamic modulation (top), and with dynamic modulation (middle) using SNR (bottom) to switch between modes of operation

UNCLASSIFIED

Consider now the scenario shown in Figure 13-11 (middle). As the SNR increases for users located towards the centre of the cell, we find that we can move to higher-order modulation schemes whilst maintaining the required BER and QoS. Thus, users located at the cell boundary use BPSK as before. Moving towards the cell centre we find that we move first to QPSK (two bits per symbol), then to 8 PSK (three bits per symbol) and finally 16QAM (four bits per symbol). What is the advantage of this? Assuming that the symbol rate is maintained, the higher order modulation schemes mean that the channel data rate is increased. Therefore, users operating QPSK can transmit the same amount of data as BPSK users in half the time, 8 PSK users can transmit the same amount of data as BPSK users in a third of the time and so on. In other words, TDMA can be used to support four 16QAM users, three 8 PSK users or two QPSK users for every BPSK user in the inner regions. Thus, overall, assuming a uniform distribution of users, system capacity can be improved significantly. Alternatively, of course, user's towards the cell's centre might use the higher-order modulation schemes simply to achieve greater data rates.

To consider how dynamic modulation might help improve capacity we will use a simple example. Table 13-2 lists the theoretical SNR required for a BER of 10^{-4} in a Gaussian channel [379]. Figure 13-12 (top) shows downlink SNR for users of the centre cell in a network of hexagonal cells assuming a reuse factor of three and a pathloss coefficient of 3.5. From Figure 13-12 (top) we can generate the cumulative distribution function for SNR within the centre cell. This is shown in Figure 13-12 (bottom left). Finally, using the SNR values listed in Table 13-2, we find the proportional distribution of the supported modulation schemes across the centre cell, Figure 13-12 (bottom right).

Modulation	Bits/Symbol	E_b/N_0	SNR
BPSK	1	8.4 dB	8.4 dB
QPSK	2	8.4 dB	11.4 dB
8-PSK	3	11.8 dB	16.6 dB
16QAM	4	12.2 dB	18.2 dB

Table 13-2: Example E_b/N_0 and SNR requirements for a $\text{BER} = 10^{-4}$ in a Gaussian channel [379]

UNCLASSIFIED

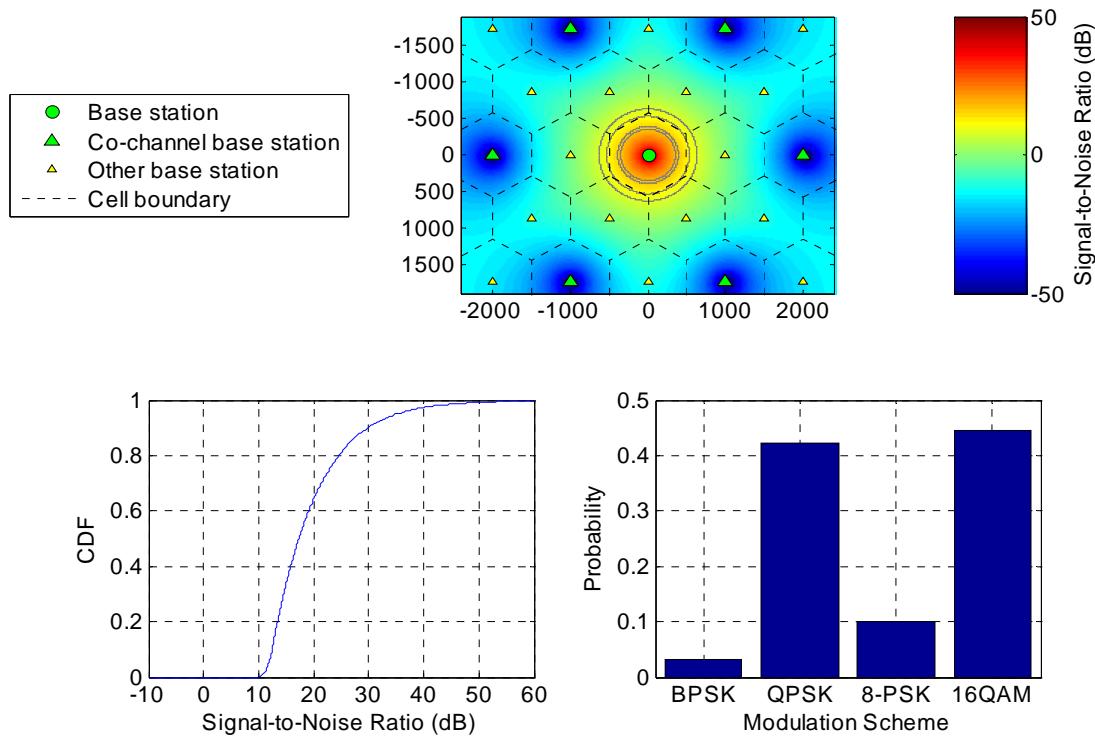


Figure 13-12: A simple example to illustrate spectral efficiency gain through dynamic modulation; SNR for a cell in a network (top); the cumulative distribution function of SNR within the centre cell (bottom left) and the proportional distribution of modulation schemes across the cell (bottom right)

Assuming a uniform distribution of users, we find the capacity of our network is, theoretically, increased almost threefold compared to the scenario using a BPSK modulation scheme only. We note that this is a simplistic example that assumes that users can be ‘packed’ into the available radio resources with 100% efficiency. Furthermore, it does not take into account additional overheads caused by an increase in control traffic. However, it clearly demonstrates that, in principle, significant capacity gains might be made through the application of dynamic modulation schemes.

In the example given above we chose to adjust the modulation method only. A similar effect might be achieved by adjusting the FEC rate. Users benefiting from higher SNR do not require the same level of FEC performance as users located at the cell boundary. By reducing the FEC coding overhead for these users additional users may be supported compared to the scenario with a static FEC scheme. Combining dynamic modulation with dynamic channel coding will permit the number of user data bits per symbol to be varied with greater granularity, thereby allowing spectral efficiency to be optimised further.

SDRs are well suited to providing the base band processing flexibility required to implement dynamic modulation and/or dynamic channel coding. Adding intelligence to enable the most appropriate waveform configuration to be selected allows a cognitive system to be realised. Note that in order to select the most appropriate waveform configuration a feedback path from the receiver to the transmitter is required to report estimated SNR data.

13.6 Conclusions

In this chapter we have considered how SDR and CR-enabled SDR might be used to improve the efficiency with which the RF spectrum is utilised within the UK. Ways of improving spectrum utilisation are needed because it will not always be practical to simply allocate previously unallocated spectrum at higher frequency ranges in response to continuing demand.

Ofcom intends to revise the way in which the spectrum in the UK is managed. First it wishes to introduce the concept of spectrum trading and move towards the use of market mechanisms to manage licensed spectrum and away from the so-called command and control method used currently. Ofcom also plans to increase the spectrum allocated to unlicensed applications to allow greater use of innovative, short-range radio technologies.

Spectrum trading will make RF spectrum a tradable asset and will provide an incentive to users to use ‘their’ spectrum as efficiently as possible. Moreover, a move towards spectrum liberalisation, i.e., the relaxation of the conditions controlling how each band of spectrum may be used, will give users the freedom to use technological advances to deploy more spectrally efficient systems. SDR and, in particular, CR-enabled SDR might prove to be key in the realisation of such systems.

The generation of realistic, qualitative values representing the efficiency gains that might be made using SDRs to implement the techniques discussed in this report requires the development of a complex simulation model. For example, a suitable model would need to model the radio environment; the movement of mobile radio terminals; and dynamic traffic loads. The development of such a comprehensive model was outside the scope of this study. However, in the absence of such a model, we have used discussion and a number of simple examples to illustrate the key concepts.

The potential of SDR to improve spectral efficiency was considered in three areas. First we considered cognitive, multi-mode terminals. Second we considered dynamic spectrum allocation and finally we considered the use of dynamic waveforms to help maximise the spectral efficiency of a particular system. Whilst we have considered these concepts in isolation, we note that they might be combined in practice to great effect. Thus, for example, we might envisage a multi-mode terminal implementing new, spectrally efficient, dynamic waveforms incorporating dynamic spectrum allocation as well as various legacy waveforms.

Multi-mode CRs have the potential to improve spectral efficiency by allowing greater utilisation to be made of short-range systems operating in unlicensed spectrum. Not only might this provide the user with a more cost-effective communications solution, but congestion in wider-area cellular networks may be relieved.

UNCLASSIFIED

DSA has the potential to deliver significant spectrum utilisation gains. Using analysis based on the Erlang B formula, we showed that allocating a finite number of channels to a single system (or, multiple systems using DSA) offers significantly greater spectral efficiency overall than allocating the same channels in a static manner to multiple systems. Different forms of DSA may be realised. For example, a central broker might be used to control access to licensed spectrum. Alternatively, devices operating in unlicensed spectrum (and perhaps selected licensed bands) might use high-performance sensing techniques to find and utilise ‘unused’ spectrum. We note that, in the latter case in particular, measures must be implemented to mitigate the probability of causing interference to existing legacy spectrum users.

Finally, dynamic waveform configuration might be used to improve the spectral efficiency of a particular radio system. For example, dynamic, closed-loop power control might be used to minimise interference caused to other co-channel users. Thus, independent radio systems sharing the same frequency band might be able to coexist with reduced separation. Alternatively, dynamic modulation and/or channel coding might be used to take advantage of high received SNR close to the transmitter to allow system capacity to be improved without requiring additional spectral resources.

We conclude that radio systems using techniques such as dynamic spectrum allocation and/or dynamic waveform control might be used to realise significant gains in the efficient utilisation of RF spectrum within the UK. SDR, through its inherent flexibility, is well suited to the implementation of these techniques. A slightly less ambitious application of radios with cognitive capabilities is that of multi-mode, ‘always best connected’ radios.

Again, SDR platforms are well suited to such applications because they may be reconfigured rapidly to switch between radio systems using common hardware resources. Furthermore, SDR-based terminals might be designed to support post-manufacture software updates to support new or updated waveforms - a feature not practical in terminals implemented using more traditional radio architectures.

14 Conclusions

The aim of this study has been to provide an evaluation of SDR with particular reference to the following topics:

- Antennas
- RF Linearisation
- Antenna Processing
- MIMO Technology and SDR
- Waveforms
- Software Aspects
- Security
- Radio Management and CR
- Regulatory Issues
- Commercial Drivers
- Assessment of SDR's First Applications and Areas of Deployment
- Spectrum Efficiency Gains of SDR and CR.

We note that some of these topics, namely, antennas, RF linearisation, MIMO, radio management and CR are not necessarily specific to SDR, or even a requirement of SDR.

However, when considering the long-term vision for SDR, i.e., the existence of highly reconfigurable, wideband, multi-mode radio platforms that can be reconfigured 'on-the-fly', a broad range of additional topics including antennas, RF linearisation, MIMO, radio management and CR become important.

Each of the main sections within this report culminates with its own set of conclusions. In addition, Section 14-1 provides a forecast for technology development epoch estimation related to SDR.

Subsequently, general conclusions are presented in Section 14-2, and finally, a summary is presented in Section 14-3.

14.1 Epoch Estimation for Technology Development

Epoch estimation for the technology development of SDR is presented in Figure 14.1. It is important to note that the estimates provided within this figure, and indeed in various sections within this study, contain a certain degree of uncertainty, as is common to any forecast.

There is a wide range of factors, from the rate of advance of research and the discovery of new technology, to the market demands for a particular requirement which will influence the rate of development of a technology, and can therefore severely distort these predictions.

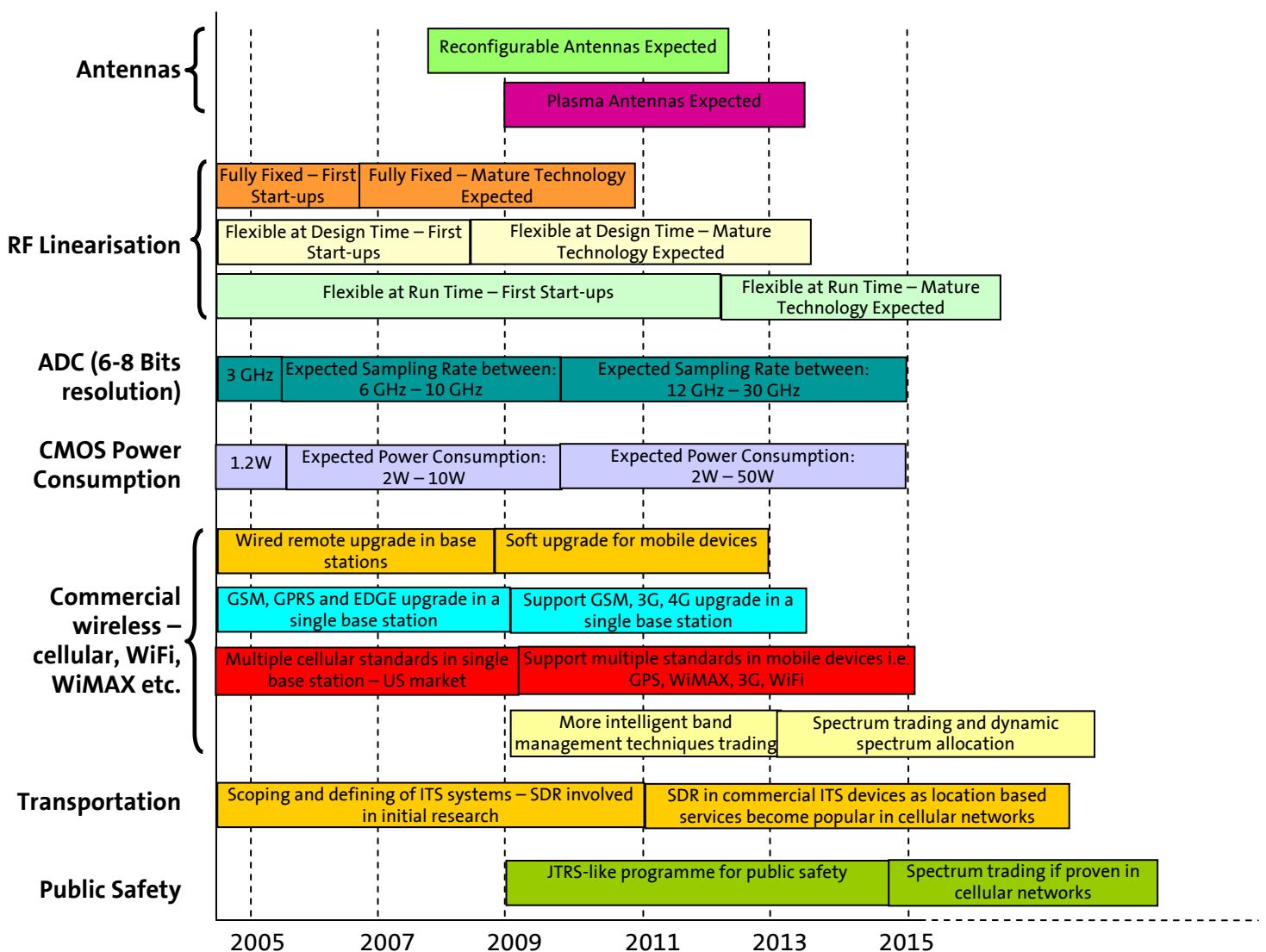


Figure 14.1: SDR-related Epoch Estimations

14.2 General Conclusions

A broad collection of topics associated with SDR have been reviewed in this study, which has culminated in a two-part document providing an evaluation of SDR. This Main Document provides a detailed evaluation of SDR, and the companion Overview Document presents a concise evaluation of SDR on a range of subjects addressed within this study, thereby ensuring an informative reference document on the subject of SDR.

Included within this evaluation of SDR are discussions on CR, a candidate technology which will intelligently manage a future SDR. The overall goal of SDR, in conjunction with the various additional technologies, such as CR, is to provide an optimal radio spectrum environment.

CR should enable signals transmitted by the radio to be modified in a manner that allows wireless communications with minimal impedance to the user of the radio device (e.g. removing issues such as the “network busy” icon possibly observed on mobile phones, or the long delays sometimes experienced when receiving wireless video imaging).

Therefore, this will allow Ofcom (and equivalent regulatory bodies in other countries) to make provision for the optimum utilisation of radio spectrum.

There are many complexities which need to be overcome to enable this ideal vision of minimal impediments for wireless communications. These range from the current limitations of the actual physical radio device through to issues of regulation.

Therefore, on the one hand there are issues on the development of technology to enable these concepts to materialise, and on the other hand, the need to ensure the regulation on radio spectrum supports dynamic capabilities of future radios.

The assessment of the various issues related to SDR presented in the two documents has been structured in a manner that allows the potential obstacles and enablers of SDR to be identified and discussed initially, which is then followed by an assessment of the first application and areas of deployment of SDR. The final part of the study provides an assessment of the spectrum efficiency gains of SDR and CR.

The various issues related to SDR, ranging from the physical implementation issues associated with SDR through to the radio management issues are considered in detail.

Key groups of investigation undertaken within this study include an assessment of antennas, RF linearisation, antenna processing, MIMO technology, waveforms and software aspects. In addition, the issues related to security, radio management, regulatory issues, commercial drivers, an assessment of SDR and spectrum efficiency gains have also been considered.

Various conclusions are raised within this pair of documents, which indicate the significance of the key technologies and how they are expected to develop over the course of the next ten, or so, years. It can be seen that the emergence of SDR is a practical, realisable technology, which marks a significant milestone in the evolution of radio.

14.3 Summary

Many topics associated with SDR have been considered in this study. Throughout the course of the report we have tried to identify potential obstacles and enablers of SDR. An informative reference document has been generated. In addition, an assessment of SDR has been provided including some consideration of the spectrum efficiency gains of SDR and CR.

It is important to recognise that SDR, as a technology, is still very much in its infancy. Consequently, it will be essential that the progress of SDR as a viable radio technology is reviewed regularly over the coming years as the technology matures and the distinction between practical reality and long-term, idealised vision becomes less blurred.

We conclude that SDR heralds the beginning of an exciting new era in the development of novel, spectrally-efficient wireless communication systems. Moreover, the development of flexible SDR platforms will represent a significant stepping-stone towards the realisation of networks of CRs and of comparable technologies.

References / Bibliography

1. Sturman, T.A. (Editor), Burr, A., Fitzpatrick, J., James, T., Rupp, M. and Weiss, S., *An Evaluation of Software Defined Radio – An Overview*, March 2006.
2. SDR Forum, “SDR Forum FAQs – How does the Forum define SDR?”,
<http://www.sdrforum.org/faq.html#define> [accessed 29 Nov 2005]
3. Joint Tactical Radio System Program, “Joint Tactical Radio System (JTRS) Operational Requirements Document (ORD)”, Version 3.2, JROCM 087-03, 9 April 2003, [downloaded from <http://jtrs.army.mil/documents/JROC%20Approved%20ORD%20v3.2%209Apr03.pdf>, 16 March 2005].
4. Multiple Access communications Limited, *Obstacles and Enablers of Software Defined Radio(SDR)*, MAC-EXFR-19-182-01.01, August 2005.
5. Fitzpatrick, J, *WP2 – Assessment of SDR’s First Applications and Areas of Deployment*, QINETIQ/05/00953 V1.0, June 2005.
6. Multiple Access communications Limited, *Spectrum Efficiency Gains of Software Defined Radio and Cognitive Radio*, MAC-EXFR-20-002-01.00, July 2005.
7. Cave M, *Review of Radio Spectrum Management: An Independent Review for Department of Trade and Industry and HM Treasury*, March 2002.
8. Ofcom, *Spectrum Framework Review*, 23 November 2004.
9. Kenington P B, “Emerging Technologies for Software Radio”, Electronics & Communication, vol 11, no 2, pp 69–83, April 1999.
10. MacLeod J R, Nesimoglu T, Beach M A and Warr P A, “Enabling Technologies for Software Defined Radio Transceivers”, Proc IEEE Military Comms Conference (MILCOM) 2002 vol 1, October 2002.
11. Padgett J E, Gunther C G and Hattori T, “Overview of Wireless Personal Communications”, IEEE Comms, vol 33, no 1, pp 28–41 January 1995.
12. Mizuno M and Ohgane T, “Application of Adaptive Array Antennas to Radio Communications”, Electronic Comms Japan, vol 77, 1994.
13. Swales A C, Beach M A, Edwards D J, and McGeehan J P, “The Performance Enhancement of Multi-Beam Adaptive Base-Station Antennas for Cellular Land Mobile Radio Systems”, IEEE Trans Veh Tech, vol 39, no 1, pp 56–67, January 1990.
14. Zetterberg P and Ottersten B, “The Spectrum Efficiency of a Base-Station Antenna Array System for Spatially Selective Transmission”, IEEE Trans Veh Tech, vol 44, no 3, pp 651–660, August 1995.
15. Harrington R F, “Effect of Antenna Size on Gain, Bandwidth and Efficiency”, Research of the NBS-D, vol 64D, no 1, January 1960.
16. Aberle J T, Oh S-H, Auckland D T and Rogers S D, “Reconfigurable Antennas for Portable Wireless Devices”, IEEE Antennas & Propagation, vol 45, no 6, pp 148–154, December 2003.

UNCLASSIFIED

17. Fathy A E, Rosen A, Owen H S, McGinty F, McGee D J, Taylor G C, Amantea R, Swain P K, Perlow S M and ElSherbiny M, "Silicon-Based Reconfigurable Antennas – Concepts, Analysis, Implementation and Feasibility", IEEE Trans Microwave Theory & Tech, vol 51, no 6, pp 1650–1661 June 2003.
18. Linden D S, "In-Situ Evolution of a Reconfigurable Antenna", IEEE Proc Aerospace Conference, vol 5, pp 2333-2338, March 2001.
19. Waldschmidt C and Wiesbeck W, "Compact Wideband Multimode Antennas for MIMO and Diversity", IEEE Trans Antennas & Propagation, vol 52, no 8, p 1963-1969 August 2004.
20. Nepa P, Serra A A, Marsico S and Manara G, "A Dual-Band Antenna for Wireless Communication Terminals", IEE Antennas and Propagation Society Symposium, vol 4, 25 June 2004.
21. Cohen N, "Fractal Antennas – part 1", Comms Quarterly, Summer edition, 1995.
22. Ansoft HFSS online user manuals, Ansoft Corporation.
23. Raab F H, Asbeck P, Cripps S, Kenington P B, Popovic Z B, Pothecary N, Sevic J F and Sokal N O, "Power Amplifiers and Transmitters for RF and Microwave", IEEE Trans Microwave Theory & Tech, vol 50, no 3, pp 814–826, March 2002.
24. SDR Forum, www.sdrforum.org [accessed 22 April 2005].
25. Parsons K J and Kenington P B, "The Efficiency of a Feedforward Amplifier with Delay Loss", IEEE Trans Veh Tech, vol 43, no 2, pp 407–412, May 1994.
26. Faulkner M, "Amplifier Linearisation Using RF Feedback and Feedforward Techniques", IEEE Trans Veh Tech, vol 47, no 1, pp 209–215, February 1998.
27. Muonen K J, Kavehrad M and Krishnamoorthy R, "Look-Up Table Techniques For Adaptive Digital Predistortion, A Development And Comparison", IEEE Trans Veh Tech, vol 49, no 5, pp 1995–2002, September 2000.
28. Kenington P B, Wilkinson R J and Parsons K J, "Noise Performance of a Cartesian Loop Transmitter", IEEE Trans Veh Tech, vol 46, no 2, pp 467–476, May 1997.
29. Eun C and Powers E J, "A New Volterra Predistorter Based on the Indirect Learning Architecture", IEEE Trans Sig Proc, vol 45, no 1, pp 223–227, January 1997.
30. Zhu A and Brazil T J, "An Adaptive Volterra Predistorter for the Linearisation of HF High-Power Amplifiers", in Proc IEEE Int Microwave Theory Tech Symp, Seattle, vol I, pp 461–464, May 2002.
31. Kenington P B, *High-linearity RF Amplifier Design*, Artech House Books, Boston/London 2000.
32. Wilkinson T A and Jones A E, "Minimisation of the Peak-to -Mean Envelope Power Ratio of Multicarrier Transmission Schemes By Block Coding", Proc 45th IEEE Veh Tech Conference (VTC 1995), Chicago, vol 2, pp 825–829, July 1995.
33. Shepherd S, Orriss J and Barton S, "Asymptotic Limits in Peak-to-Mean Envelope Power Reduction by Redundant Coding in Orthogonal Frequency-Division Multiplex Modulation", IEEE Trans Comms, vol 46, no 1, pp 5–10, January 1998.
34. Jones A E and Wilkinson T A, "Combined Coding for Error Control and Increased Robustness to System Nonlinearities in OFDM", in Proc 46th IEEE Veh Tech Conference (VTC 1996), Atlanta, vol 2, pp 904–908, 28 April – 1 May 1996.

UNCLASSIFIED

35. Tarokh V and Jafarkhani H, "On the Computation and Reduction of the Peak-to-Average Power Ratio in Multicarrier Communications", IEEE Trans Comms, vol 48, no 1, pp 37–44, January 2000.
36. Armstrong J, "Peak-to-Average Power Reduction for OFDM by Repeated Clipping and Frequency Domain Filtering", IEE Elec Lett, vol 38, no 5, pp 246–247, February 2002.
37. Li X and Cimini L J, "Effects of Clipping and Filtering on the Performance of OFDM", IEEE Comm Lett, vol 2, no 5, pp 131–133, May 1998.
38. Saul A, "Analysis of Peak Reduction in OFDM Systems Based on Recursive Clipping", in Proc 8th Int OFDM-Workshop, Hamburg, 24-25 September 2003.
39. Paterson K G, "On Codes with Low Peak-to-Average Power Ratio for Multi-Code CDMA", IEEE Trans Inf Theory, vol 50, no 3, pp 550–559, 2004.
40. Paterson K G, "Sequences for OFDM and Multi-Code CDMA: Two Problems in Algebraic Coding Theory", in Proc Sequences and their Applications - SETA01, Helleseth T, Kumar P V, Yang K (eds), Discrete Mathematics and Theoretical Computer Science Series, Springer, pp 46–71, 2002.
41. Fazel K and Kaiser S, "Analysis of Non-linear Distortions on MC-CDMA", IEEE Intl Conference on Comms (ICC 98), Atlanta, vol 2, pp 1028–1034, 7-11 June 1998.
42. Natarajan B and Nassar C, "Crest Factor Consideration in MC-CDMA with Carrier Interferometry Codes", IEEE Pacific Rim Conference Comms, Computers and Sig Proc (PACRIM 2001), Victoria, BC, Canada, vol 2, pp 445–448, Aug 2001.
43. Natarajan B, Nassar C R, Shattil S, Michelini M and Wu Z, "Application of Interferometry to MC-CDMA", accepted for IEEE Trans Veh Tech, 2005.
44. Nassar C R, Natarajan B and Shattil S, "Introduction of Carrier Interference to Spread Spectrum Multiple Access", Proc IEEE Emerging Technologies Symposium on Wireless Comms. Systems , Dallas, April 12-13 1999.
45. Choi B, Kuan E and Hanzo L, "Crest Factors of Shapiro-Rudin Sequence Based multi-code MC-CDMA signals", Proc 55th IEEE Veh Tech Conference (VTC '02), Vancouver vol 3, pp 1472–1476, 24-28 September 2002.
46. Choi B and Hanzo L, "Crest Factor of Complementary-Sequence-Based Multicode MC-CDMA Signals", IEEE Trans Wireless Comm, vol 2, no 6, pp 1114–1119, November 2003.
47. Kim J and Konstantinou K, "Digital Predistortion of Wideband Signals Based on Power Amplifier Model with Memory", IEE Elec Lett, vol 37, no 23, pp 1417–1418, November 2001.
48. Aschbacher E and Rupp M, "Identification of a Nonlinear Power Amplifier L-N-L Structure for Pre-Distortion Purposes", in Proc 14th IEEE Intl Personal, Indoor, and Mobile Radio Conference (PIMRC 2003), Beijing, vol 3, pp 2102–2106, 7-10 September 2003.
49. Boyd S, Chua L O and Desoer C A, "Analytical Foundations of Volterra Series", IMA Math Control & Inf, vol 1, pp 243–282, 1984.
50. Boyd S and Chua L O, "Fading Memory and the Problem of Approximating Nonlinear Operators with Volterra Series", IEEE Trans Circuits Syst, vol 32, no 11, pp 1150–1161, November 1985.

UNCLASSIFIED

51. Schetzen M, *The Volterra and Wiener Theories of Nonlinear Systems*, J Wiley & Sons, Chichester, 1980.
52. Saleh A A M, "Frequency-independent and frequency-dependent Nonlinear Models for TWT Amplifiers", IEEE Trans Comm, vol 29, no 11, pp 1715–1720, November 1981.
53. Kang H W, Cho Y S and Youn D H, "Adaptive Precompensation of Wiener Systems", IEEE Trans Sig Proc, vol 46, no 10, pp 2825–2829, 1998.
54. Zhou D and Debrunner V, "A Simplified Adaptive Nonlinear Predistorter for High Power Amplifiers Based on the Direct Learning Algorithm", Proc IEEE Int Conference Acoustics, Speech, and Signal Processing (ICASSP 2004), vol 4, pp 1037–1040, 17-21 May, Montreal, Canada, 2004.
55. Aschbacher E, Steinmair M and Rupp M, "Iterative Linearisation Methods Suited for Digital Pre-distortion of Power Amplifiers", in Proc 38th Asilomar Conference Signals, Systems, and Computers (ASILOMAR 2004), California, vol 2, pp 2198-2202, 7-10 November 2004.
56. Sayed A H, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, Chichester 2003.
57. Haykin S, *Adaptive Filter Theory*, Prentice Hall, 4th Ed, 2002.
58. Rupp M and Sayed A H, "Robustness of Gauss-Newton Recursive Methods: A Deterministic Feedback Analysis", Sig Proc, vol 50, pp 165–187, 1996.
59. Sayed A H and Rupp M, "Error-energy Bounds for Adaptive Gradient Algorithms", IEEE Trans Sig Proc, vol 44, pp 1982–1989, August 1996.
60. Rupp M and Sayed A H, "Supervised Learning of Perceptron and Output Feedback Dynamic Networks: A Feedback Analysis via the Small Gain Theorem", IEEE Trans Neural Net, vol 8, pp 612–622, May 1997.
61. Aschbacher E and Rupp M, "Robust Identification of an L-N-L System", Proc 37th Asilomar Conference Signals, Systems, and Computers (ASILOMAR 2003), California, vol 1, pp 1298-1302, 9-12 November 2003.
62. Nowak R D and van Veen BD, "Volterra Filter Equalization: A Fixed-Point Approach", IEEE Trans Sig Proc, vol 45, no 2, pp 377–387, February 1997.
63. Luenberger D G, *Optimization by Vector Space Methods*, John Wiley & Sons, Chichester, 1969.
64. Walden R H, "Analog-to-Digital Converter Survey and Analysis", IEEE Selected Areas of Comms, vol 17, no 4, pp 539–549, April 1999.
65. Wepman J A, "Analog-to-Digital Converters and Their Application in Radio Receivers", IEEE Comms Magazine, vol 33 no.5 pp 39–45, May 1995.
66. Analog Devices, Data sheet A/D converter AD9430, 2003.
67. Vessal F and Salama C A T, "An 8-Bit 2-Gsamples/s Folding-Interpolating Analog-to-Digital Converter in SiGe Technology", IEEE Solid-States Circuits, vol 38, no 1, pp 238–241, January 2004.
68. Uyttendhoove K and Steyaert M S J, "A 1.8V 6-bit 1.3 GHz flash ADC in 0.25- μ m CMOS", IEEE Solid-State Circuits, vol 38, no 7, pp 1115–1122, July 2003.

UNCLASSIFIED

69. Scholtens P C S and Vertregt M, "A 6-b 1.6-Gsamples/s flash A/DC in 0.18- μ m CMOS Using Averaging Termination", IEEE Solid-State Circuits, vol 37, no 12, pp 1599–1609, December 2002.
70. Geelen G, "A 6-bit 1.1 Gsamples/s CMOS A/D converter", in Proc IEEE Solid-State Circuits Conference Dig Tech Papers, pp 128–129, 2001.
71. Analog Devices, Data sheet D/A converter AD9726, 2003.
72. Van den Bosch A, Borremans M A F, Steyaert M S J and Sansen W, "A 10-bit 1-GSample/s Nyquist Current-Steering CMOS D/A Converter", IEEE Solid-State Circuits, vol 36, no 3, pp 315–324, March 2001.
73. Park S, Kim G, Park S C and Kim W, "A Digital-to-Analoge Converter Based on Differential Quad-Switching", IEEE Solid-State Circuits, vol 37, no 11, pp 1335–1338, October 2002.
74. Sandbridge Technologies, www.sandbridgetech.com, [accessed 22 April 2005].
75. Intrinsic, www.intrinsicity.com, [accessed 22 April 2005].
76. Maxfield C, "Reconfiguring Chip Design", EEdesign, September 27, 2002.
77. Brakensiek J, Oelkrug B, Bucker M, Uffmann D, Droege A, Darianian M and Otte M, "Software Radio Approach for Reconfigurable Multi Standard Radios", Proc 13th IEEE Intl Personal, Indoor, and Mobile Radio Conference (PIMRC 2002), Portugal, vol 1, pp 110–114, 15–18 September 2002.
78. Heysters P M, Bouma H, Smit J, Smit G J M and Havinga P J M, "A Reconfigurable Function Array Architecture for 3G and 4G wireless terminals", Proc World Wireless Congress, San Francisco, USA, pp 399–404, May 28–31 2002.
79. Adelante Technologies, www.adelantetech.com, [accessed 22 April 2005].
80. Tuttlebee W (ed), *Software Define Radio: Enabling Technologies*, John Wiley & Sons, Chichester, 2002.
81. Zhang H, Prabhu V, George V, Wan M, Benes M, Abnous A and Rabaey J M, "A 1-V Heterogeneous Reconfigurable DSP IC for Wireless Baseband Digital Signal Processing", IEEE Solid-State Circuits, vol 35, pp 1697–1704, November 2000.
82. Hauser J R and Wawrynek J, "GARP: a MIPS Processor with a Reconfigurable Co-processor", Proc 5th Annual IEEE Symp FPGAs for Cust Comp Machines, California, pp 12–21, April 16–18 1997.
83. PicoChip Designs Ltd, www.picochip.com, [accessed 22 April 2005].
84. Taylor M B, Kim J, Miller J, Wentzlaff D, Ghodrat F, Greenwald B, Hoffman H, Johnson P, Lee J-W, Lee W, Ma A, Saraf A, Seneski M, Shnidman N, Strumpen V, Frank M, Amarasinghe S and Agarwal A, "The RAW Microprocessor: A Computational Fabric for Software Circuits and General Purpose Programs", IEEE Micro, vol 22, no 2, pp 25–35, March/April 2002.
85. QuickSilver Technology, www.qstech.com, [accessed 22 April 2005].
86. Triscend Technologies, *White paper: Configurable Processors: An Emerging Solution for Embedded System Design*, 1998.
87. Chris F, Rennie K, Xing G, Berg S, Bolding K, Naegle J, Parshall D, Portnov D, Sulejmanasic A and Ebeling C, "An Emulator for Exploring RaPiD Configurable Computing Architectures", Proc FPL 2001, Belfast, N Ireland, pp 17–26, August 2001.

UNCLASSIFIED

88. Schmit H, Whelihan D, Tsai A, Moe M, Levine B A and Taylor R R, "PipeRench: A Virtualized Programmable Datapath in 0.18micron Technology", IEEE CICC, 2002.
89. Graham P and Nelson B, " Reconfigurable Processors for High Performance Embedded Digital Signal Processing", Proc 9th Intern Workshop FPL, August 1999.
90. Walden R H, "Performance Trends for Analog-to-Digital Converters", IEEE Comms Magazine, vol 37, no 2, pp 96–101, February 1999.
91. Brannon B, Cloninger C, Efstathiou D, Hendriks P and Zvonar Z, "Data Conversion in Software Defined Radios", in Tuttlebee W (ed), *Software Defined Radio: Enabling Technologies*, ch 4, pp 99–126, John Wiley & Sons, Chichester, 2002.
92. Beach M, Warr P and MacLeod J, "Radio Frequency Translation for Software Defined Radio", in Tuttlebee W (ed), *Software Defined Radio: Enabling Technologies*, ch 2, pp 25–78, John Wiley & Sons, Chichester, 2002.
93. Vaidyanathan P P, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, 1993.
94. Vaughan R G, Scott N L and White D R, "The Theory of Bandpass Sampling", IEEE Trans Sig Proc, vol 39, no 9, pp 1973–1984, September 1991.
95. Coulson A J, "A Generalization of Nonuniform Bandpass Sampling", IEEE Trans Sig Proc, vol 43, no 3, pp 694–704, March 1995.
96. Hsu H P, "Sampling", in Gibson J (ed), The Mobile Communications Handbook, ch 2, pp 13–22, CRC Press, Boca Raton, Florida, USA, 1996.
97. Weiss S and Stewart R W, *On Adaptive Filtering in Oversampled Subbands*, Shaker Verlag, Aachen, Germany, 1998.
98. Weiss S, Stewart R W and Davis G M, "Noise and Digital Signal Processing." in Davis GM (ed), *Handbook on Noise Reduction in Speech Applications*, ch 1, CRC Press, Cambridge, 2001.
99. Loumeau P, Navier J F, Petit H, Naviner L and Desgreys P, "Analog to Digital Conversion: Technical Aspects", Annals of Telecommunications, vol 57, no 5–6, pp 338–385, May–June 2002.
100. Bonnet C, Caire G, Enout A, Humblet P A, Montalbano A and Nussbaum D, "Open Software Radio Platform for New Generations of Mobile Communication Systems", in 3rd European DSP Education and Research Conference, Paris, France, September 2000.
101. Jian M, Bai S R, Heng K T and Yung W H, "A Software Radio Development Platform PCP200 – Partnering 'C6x with Virtex FPGA'", in 3rd European DSP Education and Research Conference, Paris, France, September 2000.
102. Weiss S, Shligersky A, Abendroth S, Reeve J S, Moreau L A V, Dodgson T E and Babb D, "A Software Defined Radio Testbed Implementation", in IEE Colloquium on DSP Enabled Radio, pp 268–274, Livingston, Scotland, September 2003.
103. Tietze U and Schenk C, *Halbleiter-Schaltungstechnik*, (10th edition), Springer Verlag, Berlin, Germany, 1993.
104. Norsworthy S R, Schreier R and Temes G C (eds), *Delta-Sigma Data Converters, Theory, Design, and Simulation*, IEEE Press, Piscataway, NJ, USA, 1997.

UNCLASSIFIED

105. Stewart R and Pfann E, "Oversampling and Sigma-Delta Strategies for Data Conversion", IEE Elect & Comms Eng, no 2, pp 37–47, February 1998.
106. Brock D K, "Superconductor Microelectronics: A Digital RF Technology for Software Radio", in Tuttlebee W (ed), Software Defined Radio: Enabling Technologies, ch 5, pp 127–150, John Wiley & Sons, Chichester, 2002.
107. Kopmann H, "Comprehensive Model-Based Error Analysis of Multiple Concurrent, Time-Interleaved, and Hybrid Ultra-Wideband Analogue-to-Digital Conversion", Sig Proc, vol 84, no 10, pp 1837–1859, October 2004.
108. Eshraghi A and Fiez T S, "A Comparative Analysis of Parallel Delta-Sigma ADC Architectures", IEEE Trans Circuits & Systems I, vol 51, no 3, pp 450–458, March 2004.
109. Elbornsson J, Gustafsson F and Eklund J E, "Blind Adaptive Equalization of Mismatch Errors in a Time-Interleaved A/D Converter System", IEEE Trans Circuits & Systems I, vol 51, no 1, pp 151–158, January 2004.
110. Juodawlkis P W, Twichell J C, Betts G E, Hargreaves J J, Younger R D, Wasserman J L, O'Donnell F J, Ray K G and Williamson R C, "Optically Sampled Analog-to-Digital Converters", IEEE Trans Microwave Theory Tech, vol 49, no 10, pp 1840–1853, October 2001.
111. Nathawad L Y, Urata R, Wooley B A and Miller D A B, "A 40-GHz-Bandwidth, 4-bit, Time-Interleaved A/D Converter Using Photoconductive Sampling", IEEE Solid-State Circuits, vol 38, no 12, pp 2021–2030, December 2003.
112. Urata R, Takahashi R, Sabnis V A, Miller D A B and Harris J S, "Ultra-Fast Optoelectronic Sample-and-Hold Using Low-Temperature-Grown GaAs MSM", IEEE Photonics Tech Letters, vol 15, no 5, pp 724–726, May 2003.
113. Han Y and Jalali B, "Photonic Time-Stretched Analog-to-Digital Converter: Fundamental Concepts and Practical Considerations", Lightwave Tech, vol 21, no 12, pp 3085–3103, December 2003.
114. Urata R, Nathawad L Y, Takahashi R, Ma K, Miller D A B, Wooley B A and Harris J S, "Photonic A/D Conversion Using Low-Temperature-Grown GaAs MSM Switches Integrated with Si-CMOS", Lightwave Tech, vol 21, no 12, pp 3104–3115, December 2003.
115. Brock D K, Mukhanov O A and Rosa A, "Superconductor Digital RF Development for Software Radio", IEEE Comms Magazine, vol 39, no 2, pp 174–179, February 2001.
116. Mukhanov O A, Gupta D, Kadin A M and Semenov V, "Superconductor Analog-to-Digital Converters", Proc of the IEEE, vol 92, no 10, pp 1564–1584, October 2004.
117. Foschini G J, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when Using Multiple Antennas", Bell Labs, vol 1, no 2, pp 41–59, Autumn 1996.
118. Foschini G J and Gans M J, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas", Wireless Personal Comms, vol 6, pp 311–335, 1998.
119. Lucent Bell Labs, "Bell Labs Scientists Shatter Limit on Fixed Wireless Transmission" [press release], 9th September 1998 [downloaded from <http://www.lucent.com/press/0998/980909.bla.html> 22 April 2005].

UNCLASSIFIED

120. Shannon C E, "A Mathematical Theory of Communication", Bell System, vol 27, pp 379–423 and 623–656, July and October, 1948.
121. Telatar E, "Capacity of multi-antenna Gaussian channels", internal AT&T technical note, 1995, later published in European Trans Telecomms, vol 10, no 6, pp 585–595, Nov–Dec 1999.
122. Golden G D, Foschini G J, Valenzuela R A and Wolniansky P W, "Detection Algorithm and Initial Laboratory Results using V-BLAST Space-Time Communication Architecture", IEE Elect Lett, vol 35, no 1, pp 14–16, January 7, 1999.
123. Tarokh V, Seshadri N and Calderbank A R, "Space-time Codes for High Data Rate Wireless Communication: Performance Criteria and Code Construction", IEEE Trans Inf Theory, vol 44, no 2, March 1998.
124. Tarokh V, Jafarkhani H and Calderbank A R, "Space-time Block Codes from Orthogonal Designs", IEEE Trans Inf Theory, vol 45 no 5, pp 1456–1467, July 1999.
125. Burr A G, "Space-time Signal Processing for Real-World Channels", Euro Microwave Conference, Amsterdam, 13th September 2004.
126. Burr A G, "Capacity Bounds and Estimates for the Finite Scatterers MIMO Wireless Channel", IEEE Sel Areas Comms, vol 21, no 5, pp 812–818, June, 2003.
127. Becher R, Dillinger M, Haardt M and Mohr W, "TI Broad-band Wireless Access and Future Communication Networks", Proc of the IEEE, vol 89, no 1, pp 58–75, January 2001.
128. Morawski R, Le-Ngoc T and Naeem O, "Wireless and Wireline MIMO Testbed", Oliver G, Pierre S, Sood VK (eds), IEEE CCECE-2003 Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Tech, cat no 03CH37436, vol 3, pp 1913–16, May 2003.
129. Araki K, "RF Analog Smart Devices/Circuits and their Applications", Trans Inst Elect, Inf and Comm Engs, vol J87-C, no 1, pp 3–11, January 2004.
130. Gifford S, Kleider J E and Chuprun S, "OFDM-MIMO Communication Systems in a Rayleigh Faded Environment with Imperfect Channel Estimates", IEEE Military Comms Conference (MILCOM) 2003, IEEE cat no 03CH37500, vol 1, pp 633–7, 2003.
131. Rao R M and Daneshrad B, "I/Q Mismatch Cancellation for MIMO-OFDM Systems", Proc IEEE 15th Intl Symp Personal, Indoor and Mobile Radio Comms, IEEE cat no 04TH8754, vol 4, pp 2710–14, 2004.
132. Stüber G L, Barry J R, McLaughlin SW and Li Y, ingram MA, Pratt TG, "Broadband MIMO-OFDM Wireless Communications", Proc of the IEEE, vol 92, no 2, pp 271–94, February 2004.
133. Eireiner T, Muller T, Luy J F and Owens F, "Implementation of a Smart Antenna System with an Improved NCMA Algorithm", 2003 IEEE MTT S Intl Microwave Symposium Digest, cat no 03CH37411, vol 3, pp 1529–32, 2003.
134. 3GPP, www.3gpp.org, [accessed 22 April 2005].
135. IEEE 802 Wireless World, <http://802wirelessworld.com>, [accessed 22 April 2005].

UNCLASSIFIED

136. 3GPP “Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD)” 3GPP specification 3G TS 25.211(V3.5.0).
137. 3GPP “Physical Layer Procedures (FDD)”, 3GPP Specification 3G TS 25.214(V3.5.0).
138. Alamouti S M, “A Simple Transmit Diversity Technique for Wireless Communications” IEEE Sel Areas Comms, vol 16, no 8, pp 1451–8, October 1998.
139. 3GPP “Tx Diversity Solutions for Multiple Antennas” 3GPP specification 3G TR25.869(V1.2.0), August 2003.
140. 3GPP “Multiple-Input Multiple Output in UTRA” 3GPP report TR 25.876(V1.7.0), August 2004.
141. Mujtaba S A, et al., “TGN Sync Complete Proposal” IEEE 802.11-04/0888r4, January 2005.
142. Jones V K, et al., “WWiSE IEEE 802.11n Proposal” IEEE 802.11-04/0935r3, September 2004.
143. FLOWS, “Design and Performance of Antenna Prototypes”, FLOWS Report D16, November 2004 [downloaded from <http://www.flows-ist.org/main/outputs/list.htm>, 31 January 2005].
144. Burr A G, “On the Channel Autocorrelation Matrix of a MIMO System”, COST 273 TD(04)108, 10th COST 273 MCM, Gothenburg, Sweden, June 2004.
145. Butler J and Lowe R, “Beam Forming Matrix Simplifies Design of Electronically Scanned Antennas”, Electron Design, vol 9, pp 170–173, April 1961.
146. FLOWS, “Development and Simulation of a Single Standard MIMO Transceiver”, FLOWS Report D9, June 2003 [downloaded from <http://www.flows-ist.org/main/outputs/D9.htm>, 31 January 2005].
147. Evans D, Khatri B and Raynes D, “Simplified MIMO Receiver using Orthogonally Coded Signals”, UK Patent Application no 0208214.7, Applicant: Koninklijke Philips Electronics NV, 10 April 2002.
148. Price R and Green P E, “A Communication Technique for Multipath Channels”, Proc of the IRE, vol 46, pp 555–570, March 1958.
149. Shen J and Burr A G, “Turbo Multiuser Receiver for Space-Time Turbo Coded Uplink CDMA over Frequency-Selective Fading Channel”, European Personal and Mobile Communications Conference (EPMCC’03), Glasgow, April 2003.
150. FLOWS, “Outline Design for Terminal Baseband Processing and Implementation Complexity”, FLOWS Report D18, November 2004 [downloaded from <http://www.flows-ist.org/main/outputs/list.htm>, 31 January 2005].
151. Torrance J M and Hanzo L, “Upper Bound Performance of Adaptive Modulation in a Slow Rayleigh Fading Channel”, IEE Elect Lett, vol 32, no 8, pp 718–719, April 11 1996.
152. Burr A G, “Bounds and estimates of the Uplink Capacity of Cellular Systems”, Proc IEEE Veh Tech Conference (VTC 1994), Stockholm, vol 3, pp 1480–4, 8–10 June 1994.

UNCLASSIFIED

153. Zhang L, Burr A G and Pearce D A J, "Capacity of Cellular System for the Finite Scatterers MIMO Wireless Channel", Wireless Personal Multimedia Conference 2004, Abano Terme, Italy, September 2004.
154. Shannon C E, "Communication in the Presence of Noise", Proc IRE, vol 37, pp 10–21, 1949.
155. Molisch A F, "MIMO Systems with Antenna Selection: An Overview", IEEE Radio & Wireless Conference (RAWCON), pp 167–170, August 2003.
156. 3GPP, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; High Speed Downlink Packet Access (HSDPA)"; Overall Description, 2004.
157. DVB Project, "Digital Video Broadcasting System for Handheld Terminals", June 2004.
158. ETSI, "Digital Cellular Telecommunications System (Phase 2+); Radio Transmission and Reception, 1999 edition", Sophia-Antipolis Cedex, France, 1999.
159. ETSI, "Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television", v1.5.1 edition, Sophia-Antipolis Cedex, France, 2004.
160. Garg V K and Wilkes J E, *Principles and Applications of GSM*, Prentice Hall, New Jersey, 1999.
161. Holma H and Toskala A (eds), *WCDMA for UMTS*, John Wiley & Sons, Chichester, October 2000.
162. IEEE, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Std 802.11, March 1999.
163. IEEE, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Std 802.11a, September 1999
164. IEEE, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, Std 802.11b, September 1999.
165. IEEE, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band*, Std 802.11g, June 2003.
166. Hannikainen M, Hamalainen T D, Niemi M and Saarinen J, "Trends in Personal Wireless Data Communications", Computer Comms, vol 25, no 1, pp 84–99, January 2002.
167. IEEE, *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Personal Area Networks (WPANs)*, Std 802.15.1, June 2002.
168. IEEE, *Wireless Medium Access Control (MAC) and Physical (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs)*, Std 802.15.3, June 2003.
169. IEEE, *Wireless Medium Access Control (MAC) and Physical (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)*, Std 802.15.4, 2003.
170. IEEE, *Air Interface for Fixed Broadband Wireless Access Systems*, Std 802.16, December 2001.

UNCLASSIFIED

171. IEEE, *Air Interface for Fixed Broadband Wireless Access Systems-Amendment 2: Medium Access Control Modifications and Additional Physical Layer Specifications*, , Std 802.16a, January 2003.
172. Sklar B, *Digital Communications Fundamentals and Applications*, Prentice Hall, New Jersey, 1988.
173. Hanzo L, Webb W and Keller T, *Single-and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting*, (1st edition), John Wiley & Sons, Chichester, England, 2000.
174. Hanzo L, Ng SX, Webb W and Keller T, *Quadrature Amplitude Modulation From Basics to Adaptive Trellis-Coded, Turbo-EQUALISED and Space-Time Coded OFDM, CDMA and MC-CDMA*, (2nd edition), John Wiley & Sons, Chichester, 2004.
175. Weinstein S B and Ebert P M, "Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform", IEEE Trans Comms, vol 19, no 5, pp 628–634, October 1971.
176. Proakis J G, *Digital Communications*, (3rd edition), McGraw-Hill, New York, 1995
177. Rappaport T S, *Wireless Communications: Principles and Practice*, Prentice Hall, New Jersey, July 1999.
178. Hanzo L, Cheriman P and Streit J, *Wireless Video Communications: Second to Third Generation and Beyond*, (1st edition) Wiley-IEEE Press, Piscataway, New Jersey, USA, February 2001.
179. Steele R and Hanzo L, *Mobile Communications*, (2nd edition) John Wiley & Sons, Chichester, 1999.
180. Parssinen A, Jussila J, Ryynanen J, Sumanen L and Halonen K A I, "A 2-GHz Wide-Band Direct Conversion Receiver for W CDMA Applications", IEEE Solid-State Circuits, vol 34, no 12, pp 1893–1903, December 1999.
181. Hull C D, Tham J L and Chu R R, "A Direct-Conversion Receiver for 900 MHz (ISM Band) Spread-Spectrum Digital Cordless Telephone", IEEE Solid-State Circuits, vol 31, no 12, pp 1955–1963, December 1996.
182. Rofougaran A, Chang G, Rael J J, Chang JY-C, Rofougaran M, Chang P J, Djafari M, Ku M-K, Min J, Roth E W, Abidi A A and Samueli H, "A Single-Chip 900 MHz Spread-Spectrum Wireless Transceiver in 1 μ m CMOS – Parts I & II", IEEE Solid-State Circuits, vol 33, no 4, pp 515–547, April 1996.
183. Caias P, Staraj R, Kossiavas G and Luxey C, "Compact Internal Multi-band Antenna for Mobile Phone and WLAN Standards", IEE Elect Lett, vol 40, no 15, pp 920–921, July 2004.
184. Karimi HR, Anderson N W and McAndrew P, "Digital Signal Processing Aspects of Software Definable Radios", in IEE Colloquium on Adaptable and Multistandard Mobile Radio Terminals, vol 3, pp 1–8, London, March 1998.
185. Karimi H R and Friedrichs B, "Wideband Digital Receivers for Multi-Standard Software Radios", in IEE Colloquium on Adaptable and Multistandard Mobile Radio Terminals, vol 5, pp 1–7, London, March 1998.
186. Baruffa G, Caeopardi S and Soazzi S M, "FPGA Implementation of a Multimodal Sample Rate Converter and Synchronizer", in Proc 13th IEEE Intl Symposium Personal, Indoor and Mobile Radio Comms, vol 1, pp 447–451, Lisbon, September 2002.

UNCLASSIFIED

187. Tecpanecatl-Xihuitl J L, Aguilar-Ponce R M, Bayoumi M A and Zavidovique B B, "Digital IF Decimation Filters for 3G Systems using Pipeline/Interleaving Architecture", in Proc Intl Conference on Sig Proc and Applications, vol 2, pp 327–330, Paris, July 2003.
188. Li W, Liu J, Wang J, Zhang C and Guo W, "An Efficient Digital IF Down-converter for Dual-mode W CDMA/EDGE Receiver Based on Software Radio", in IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication, vol 2, pp 713–716, Shanghai, June 2004.
189. Jian M, Yung W H and Songrong B, "An Efficient IF Architecture for Dual-Mode GSM/W CDMA Receiver of a Software Radio", in Proc IEEE Intl Workshop on Mobile Multimedia Comms, pp 21–24, USA, November 1999.
190. Yung W H, Jian M and Ho Y W, "Polyphase Decomposition Channelizers for Software Radios", in Proc IEEE Intl Symposium on Circuits and Systems, vol 2, pp 353–356, Geneva, May 2000.
191. Vinod A P, Lai EM-K, Premkumar A B and Lau C T, "A Reconfigurable Multi-Standard Channelizer using QMF Trees for Software Radio Receivers", in Proc 14th IEEE Intl Symposium Personal, Indoor and Mobile Radio Comms, vol 1, pp 119–123, Beijing, September 2003.
192. Efstathiou D, "SoftCellTM: a Multi-Carrier Transceiver Solution for Multi-mode Base-stations", in Proc 11th IEEE Intl Symposium Personal, Indoor and Mobile Radio Comms, vol 1, pp 469–473, London, September 2000.
193. Efstathiou D, "SoftCellTM: a Multi-Carrier Chip-set for Software Definable Radio Base-stations", in Proc Global Telecommunications Conference, vol 1, pp 172–176, San Francisco, California, December 2000.
194. Kountouris A A, Moy C, Rambaud L and Corre P L, "A Reconfigurable Radio Case Study: A Software Based Multi-Standard Transceiver for UMTS, GSM, EDGE and Bluetooth", in Proc IEEE Veh Tech Conference, vol 2, pp 1196–1200, Atlantic City, New Jersey, October 2001.
195. Kountouris A A, Moy C, Rambaud L and Corre P L, "A Software Radio Approach for the Transceiver Transition from 2G to 2.5G to 3G", in International Symposium of Sig Proc and its Applications, vol 2, pp 485–488, Kuala-Lampur, August 2001.
196. Kim M and Lee S, "Design of Dual Mode Digital Down Converter for W CDMA and cdma2000", ETRI Journal, vol 26, no 6, pp 555–559, December 2004.
197. Wiesler A, Machauer R and Jondral F, "Comparison of GMSK and Linear Approximated GMSK for use in Software Radio", in IEEE International Symposium on Spread Spectrum Techniques and Applications, vol 2, pp 557–560, Sun City, South Africa, September 1998.
198. Wiesler A and Jondral F, "Software Radio Structure for Second Generation Mobile Communication Systems", in Proc Veh Tech Conference Spring, vol 3, pp 2363–2367, May 1998.
199. Wiesler A, Schober H, Machauer R and Jondral F, "Software Radio Structure for UMTS and Second Generation Mobile Communication Systems", in Proc Veh Tech Conference Fall, vol 2, pp 939–942, Houston, Texas, September 1999.

UNCLASSIFIED

200. Jondral F, Wiesler A and Machauer R, "A Software Defined Radio Structure for 2nd and 3rd Generation Mobile Communications Standards", in Proc IEEE International Symposium on Spread Spectrum Techniques and Applications, vol 2, pp 637–640, Parsippany, USA, September 2000.
201. Wiesler A and Jondral FK, "A Software Radio for Second- and Third-Generation Mobile Systems", IEEE Trans Veh Tech, vol 51, no 4, pp 738–748, July 2002.
202. Rhiemeier A-R, *Modular Software Defined Radio*, University of Karlsruhe, Karlsruhe, 2004.
203. Kammerer K D, *Nachrichtenübertragung*, B G Teubner, Stuttgart, Germany, 1992.
204. Lee S P and Lee L T, "Free Vibration Analysis of Rectangular Plates with Interior Point Supports", Mechanics of Structures and Machines, vol 22, no 4, pp 505–538, Nov 1994.
205. Sheikh F and Masud S, "DSP Implementation of Concurrent GSM and CDMA Modems for Software Defined Radios", in Proc International Conference on Information, Communications and Sig Proc, vol 3, pp 1732–1736, Singapore, December 2003.
206. Jondral F, "Parametrization – a Technique for SDR Implementation", in Tuttlebee W (ed), *Software Defined Radio: Enabling Technologies*, chapter 8, pp 233–256, John Wiley& Sons, Chichester, 2002.
207. Ryynanen J, Kivekas K, Jussila J, Parssinen A and Halonen K A I, "A Dual-Band RF Front-End for W CDMA and GSM Applications", IEEE Journal of Solid-State Circuits, vol 36, no 8, pp 1198–1204, August 2001.
208. Ryynanen J, Kivekas K, Jussila J, Sumanen L, Parssinen A and Halonen K A I, "A Single-Chip Multimode Receiver for GSM900, DCS1800, PCS1900, and W CDMA", IEEE Journal of Solid-State Circuits, vol 38, no 4, pp 594–602, April 2003.
209. Richter R and Jentschel J H, "An Integrated Wideband-IF-Receiver Architecture for Mobile Terminals", in Proc IEEE MTT-S International Microwave Symposium, vol 1, pp A69–A72, Philadelphia, USA, June 2003.
210. Wu, Razavi B, "A 900 MHz / 1.8 GHz CMOS Receiver For Dual-Band Applications", IEEE Solid-State Circuits, vol 33, no 12, pp 2178–2185, December 1998.
211. Jung J H and Lyu D S, "An Architecture of Reconfigurable Transceiver Based on Digital IF for W CDMA and IS-95 Base Stations", in International Symposium on Wireless Personal Multimedia Communications, vol 2, pp 831–834, Honolulu, Hawaii, October 2002.
212. Kang B-G, "Design and Implementations of a Software Defined Radio Scheme Based Modem for High-Speed Multimedia Data Service", in Proc IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering, pp 889–892, Beijing, China, October 2002.
213. Lackey R I and Upmal D W, "SPEAKeasy: The Military Software Radio", IEEE Comms Magazine, vol 33, no 5, pp 56–61, May 1995.
214. Cook P G and Bonser W, "Architectural Overview of the SPEAKeasy System", IEEE Selected Areas in Comms, vol 17, no 4, pp 650–661, April 1999.

UNCLASSIFIED

215. Fox P W, "Software Defined Radios-Motorola's Wireless Information Transfer System (WITS)", in EUROCOMM Information Systems for Enhanced Public Safety and Security, pp 43–46, Munich, May 2000.
216. Willink E D, "The Waveform Description Language", in Tuttlebee W (ed), Software Defined Radio: Enabling Technologies, chapter 13, pp 365–397, John Wiley & Sons, Chichester, 2002.
217. Glossner J, Iancu D, Lu L, Hokenek E and Moudgil M, "A Software-Defined Communications Baseband Design", IEEE Comms Magazine, vol 41, no 1, pp 120–128, January 2003.
218. Helmschmidt J, Schuler E, Rao P, Rossi S, di Matteo S and Bonitz R, "Reconfigurable Signal Processing in Wireless Terminals", in Proc Design, Automation and Test in Europe Conference and Exhibition, pp 244–249, Munich, March 2003.
219. Uehara K, Shiba H, Shono T, Shirato Y, Yoshioka H, Nakatsugawa M, Kubota S and Umehira M, "Design and Performance Evaluation of Software Defined Radio Prototype for PHS and IEEE 802.11 wireless LAN", in Proc 13th IEEE Intl Symposium Personal, Indoor and Mobile Radio Comms, pp 452–457, Lisbon, September 2002.
220. Hotti M, Kaukovuori J, Ryynanen J, Kivekas K, Jussila J and Halonen K, "A Direct Conversion RF Front-End for 2-GHz W CDMA and 5.8-GHz WLAN Applications", in Radio Frequency Integrated Circuits Symposium, pp 45–48, June 2003.
221. Ratni M, Kruzevic D, Wang Z and Jurgensen J-U, "Broadband Digital Direct Down Conversion Receiver Suitable for Software Defined Radio", in Proc 13th IEEE Intl Symposium Personal, Indoor and Mobile Radio Comms, vol 1, pp 100–104, Lisbon, Portugal, September 2002.
222. Patel M and Darwazeh I, "Investigation of the Performance of a Multimode, Multi-band Receiver for OFDM and Cellular Systems", in Proc IEEE Veh Tech Conference, vol 1, pp 284–288, Orlando, Florida, USA, October 2003.
223. Uehara K, Tanaka H, Shiba H, Suzuki Y, Nakatsugawa M, Shirato Y and Kubota S, "Design of Software Radio for Cellular Communication Systems and Wireless LANs", in Proc 11th IEEE Intl Symposium Personal, Indoor and Mobile Radio Comms, vol 1, pp 474–478, London, September 2000.
224. Shono T, Shiba H, Shirato Y, Uehara K, Araki K and Umehira M, "Performance of IEEE 802.11 Wireless LAN Implemented on Software Defined Radio with Hybrid Programmable Architecture", in Proc Intl Conference on Comms, Alaska, USA, pp 2035–2040, May 2003.
225. Tibenderana C, Dodgson T E, Weiss S and Babb D, "Towards Software Defined Radio (SDR) Bluetooth and IEEE802.11b Modem Integration", in 9th Wireless World Research Forum (WWRF) Meeting, Zurich, July 2003.
226. Tibenderana C and Weiss S, "A Low-Complexity High-Performance Bluetooth Receiver", in IEE Colloquium on DSP Enabled Radio, pp 426–435, Livingston, Scotland, September 2003.
227. Tibenderana C and Weiss S, "Low-Complexity High-Performance GFSK Receiver With Carrier Frequency Offset Correction", in Proc IEEE International Conference on Acoustics, Speech and Signal Processing, vol IV, pp 933–936, Montreal, Canada, May 2004.

UNCLASSIFIED

228. Tibenderana C and Weiss S, "A Low-Cost Scalable Matched Filter Bank Receiver for GFSK Signals with Carrier Frequency and Modulation Index Offset Compensation", in ASILOMAR Conference on Signals, Systems, and Computers, Pacific Grove, California, November 2004.
229. Tibenderana C and Weiss S, "Efficient and Robust Detection of GFSK Signals under Dispersive Channel, Modulation Index and Carrier Frequency Offset Conditions", Special Issue on DSP Enabled Radio, EURASIP Appl Sig Proc, Accepted for Publication, January 2005.
230. Schiphorst R, Hoeksema F W and Slump C H, "A Bluetooth-Enabled HiperLAN/2 Receiver", in Proc IEEE Veh Tech Conference, vol 5, pp 3443–3447, Orlando, Florida, USA, October 2003.
231. Arkesteijn V J, Schiphorst R, Hoeksema F W, Nauta B and Slump C H, "A Software Defined Radio Test-bed for WLAN Front Ends", in Proc Program for Research on Embedded Systems and Software Workshop, Utrecht, Netherlands, October 2002.
232. Schiphorst R, Hoeksema F W and Slump C H, "A Flexible WLAN Receiver", in Proc Program for Research on Integrated Systems and Circuits Workshop, pp 555–562, Veldhoven, Netherlands, November 2003.
233. Schiphorst R, Hoeksema F W and Slump C H, "A (Simplified) Bluetooth Maximum A Posteriori Probability (MAP) Receiver", in Proc IEEE Workshop on Sig Proc Advances in Wireless Communications, Rome, June 2003.
234. Rauwerda G K, Heysters P M and Smit G J M, "Mapping Wireless Communication Algorithms Onto a Reconfigurable Architecture", Supercomputing, vol 30, no 3, pp 263–282, December 2004.
235. Schiphorst R, Hoeksema F W and Slump C H, "A Real-Time GPP Software-Defined Radio Testbed for the Physical Layer of Wireless Standards", EURASIP Appl Sig Proc, (to appear).
236. Zannoth M, Ruhlicke T and Klepser B U, "A Highly Integrated Dual-Band Multimode Wireless LAN Transceiver", IEEE Solid-State Circuits, vol 39, no 7, pp 1191–1195, July 2004.
237. Chen K-C and Wu S-T, "A Programmable Architecture for OFDM-CDMA", IEEE Comms Magazine, vol 37, no 11, pp 76–82, November 1999.
238. Fu P-W and Chen K-C, "A Programmable Transceiver Structure of Multi-Rate OFDM-CDMA for Wireless Multimedia Communications", in Proc IEEE Veh Tech Conference, vol 3, pp 1942–1946, Island of Rhodes, Greece, May 2001.
239. Wu X, Jiang B and Yin Q, "Software-Defined Radio Based Baseband Discrete Model for Orthogonal Frequency Division Multiplexing CDMA Systems", in Proc IEEE Veh Tech Conference, vol 2, pp 771–775, Island of Rhodes, Greece, May 2001.
240. Wu X, Zhao Z and Yin Q, "General Baseband Digital Model and Matrix Representation for Different orthogonal frequency division multiplexing CDMA schemes", in Proc Intl Conference on Comms, vol 2, pp 427–430, Helsinki, Finland, June 2001.
241. Yang L-L and Hanzo L, "Software-Defined-Radio-Assisted Adaptive Broadband Frequency Hopping Multicarrier DS-CDMA", IEEE Comms Magazine, vol 40, no 3, pp 174–183, March 2002.

UNCLASSIFIED

242. Rykaczewski P, Martoyo H, Liu Z and Jondral F K, "Multimode Detector and I/Q Imbalance Compensator in a Software Defined Radio", in Proc IEEE Radio and Wireless Conference, pp 521–524, Atlanta, USA, September 2004.
243. Petri I and Nastooth A, "An Overview of Software Defined Radio Technologies", Technical Report no 652, Turku Centre for Computer Science Finland, December 2004 ISBN: 952-12-1486-4 [downloaded from <http://www.tucs.fi/publications/attachment.php?fname=TR652.pdf>, 22 April 2005].
244. Hayes N, "Software Communications Architecture", in OMG's Software Based Communications Workshop 2004, 2004 [downloaded from http://www.omg.org/news/meetings/workshops/SBC_2004_Manual/T1_Hayes_SCAOverview.pdf, 22 April 2005].
245. Bose V, *Design and Implementation of Software Radios Using a General Purpose Processor*, PhD thesis Massachusetts Institute of Technology, June 1999 [downloaded from <http://sigmobile.org/phd/1999/theses/bose.pdf>, 22 April 2005].
246. Software-Based Communication Domain Task Force, website, <http://sbc.omg.org>, [accessed 22 April 2005].
247. Barbeau M, Bordeleau F and Smith J, "An Introduction to a UML Platform Independent Model of a Software Radio", International Conference on Telecommunications (ICT), 2002 [downloaded from <http://www.scs.carleton.ca/~barbeau/Publications/2002/TS-5-Paper.pdf>, 22 April 2005].
248. Communications Research Centre Canada, "Software Defined Radio" [online], <http://www.crc.ca/en/html/crc/home/research/satcom/rars/sdr/sdr>, [accessed 22 April 2005].
249. Koushanfar F, Potkonjak M and Sangiovanni-Vincentelli A, "Fault-Tolerance in Sensor Networks", in Mahgoub I and Ilyas M (eds), Handbook of Sensor Networks, CRC press, Section VIII, no 36, 2004 [downloaded from http://www-cad.eecs.berkeley.edu/~farinaz/Papers/chapter-FT_04.pdf, 22 April 2005].
250. The Berkeley/Stanford Recovery-Oriented Computing (ROC) Project, <http://roc.cs.berkeley.edu>, [accessed 22 April 2005].
251. W3C, "About the World Wide Web Consortium (W3C)" [online], [downloaded from <http://www.w3.org/2003/01/Consortium.pdf>, 22 April 2005].
252. W3C, "W3C Semantic Web" [online], <http://www.w3c.org/2001/sw>, [accessed 22 April 2005].
253. Berners-Lee T, Hendler J and Lassila O, "The Semantic Web", Scientific American, May 2001.
254. W3C, "Web Services Architecture" [online], <http://www.w3.org/TR/ws-arch>, [accessed 22 April 2005].
255. Farkas K, Ruf L, May M and Plattner B, "Framework For Service Provisioning In Mobile Ad Hoc Networks", International Conference on Telecommunications and Computer Networks, http://www.tik.ee.ethz.ch/~farkas/publications/Siramon-iadat_tcn2004.pdf, 2004.

UNCLASSIFIED

256. Cokus M and Winkowski D, "XML Sizing and Compression Study for Military Wireless Data", in XML Conference and Exposition, Baltimore, USA, http://www.idealliance.org/papers/xml02/dx_xml02/papers/06-02-04/06-02-04.pdf, 2002.
257. Fox G and Walker D, "e-Science Gap Analysis", Technical Report, Indiana University and Cardiff University, UK e-Science Centre, <http://grids.ucs.indiana.edu/ptliupages/publications/GapAnalysis30June03v2.pdf>, 2003.
258. Foster I, Kesselman C and Tuecke S, "The Anatomy of the Grid Enabling Scalable Virtual Organizations", International Journal of Supercomputer Applications, <http://www-fp.mcs.anl.gov/~foster>, 2001.
259. McKnight L, Howison J and Bradner S, "Wireless Grids: Distributed Resource Sharing by Mobile, Nomadic, and Fixed Devices", IEEE Internet Computing, pp 24–31, <http://dsonline.computer.org/0407/f/w4geip.htm>, July/August, 2004.
260. Bizer C and Oldakowski R, "Using Context- and Content-Based Trust Policies on the Semantic Web", poster in 13th World Wide Web Conference <http://www.wiwiss.fu-berlin.de/suhl/bizer/SWTSGuide/p747-bizer.pdf>, 2004.
261. Wang J, Kokar M M, Baclawski K and Brady D, "Achieving Self-Awareness of SDR Nodes through Ontology-Based Reasoning and Reflection", Software Defined Radio Technical Conference, Phoenix, Arizona, paper no 2.4-3, http://ucsu.colorado.edu/~weingatb/SDR_1/2.4.3kokar.pdf, October 2004.
262. Gaynor M, McKnight L, Hwang J, Freedman J, "Wireless Grid Networks and Virtual Markets", http://wirelessgrids.net/docs/CCCT03_Gaynor_T244WL.pdf, 2003.
263. Internet Engineering Task Force, "RFC 2828 - Internet Security Glossary", May 2000.
264. Internet Engineering Task Force , "RFC 1321 - MD5 Message Digest Algorithm", April 1992.
265. NIST, FIPS-180-1, "Secure Hash Standard", April 1995.
266. Cook P G, "SDR System Security", SDRF-02-A-0006-V0.00, www.sdrforum.org, November 2002.
267. SDR Forum, "Wireless Software Download Security", SDRF-04-I-0069-V0.00, www.sdrforum.org, June 2004.
268. SDR Forum, "Security Considerations for Operational Software for Software Defined Radio Devices in A Commercial Wireless Domain", SDRF-04-A-0010-V0.00, www.sdrforum.org, October 2004.
269. Cook PG, "A Structure for Software Defined Radio Security", SDRF-03-I-0010-V0.0, www.sdrforum.org, May 2003.
270. Delio M, www.wired.com/news/wireless/0,1382,44545,00.html, June 2001.
271. Express Logic, "Efficient Memory Protection for Embedded Systems" <http://www.rtcmagazine.com/home/article.php?id=100120>, September 2004.
272. XG Working Group, "The XG Vision – RFC 2.0", www.darpa.mil/ato/programs/xg/rfc_vision.pdf, 2004.

UNCLASSIFIED

273. Notor J, "Cognitive radio: an opportunity lost?", http://www.commsdesign.com/design_corner/showArticle.jhtml?articleID=29100644, 2004.
274. Leaves P, Ghaheri-Niri S, Tafazolli R, Christodoulides L, Sammut T, Stah W and Huschke J, "Dynamic Spectrum Allocation in a Multi-Radio Environment : Concept and Algorithm", *IEE Second International Conference on 3G Mobile Communication Technologies*, London, pp. 53-57, 26-28, www.ist-drive.org/papers/iee-3g2001-dsa-paper.pdf, 2001.
275. Tsui E, "What are Adaptive, Cognitive Radios?" www.bwrc.eecs.berkeley.edu/Presentations/Retreats/Winter Retreat 2004/Tuesday%20AM/ACR%20E%20Tsui.ppt, 2004.
276. Mitola J, "Cognitive Radio for Flexible Mobile Multimedia Communications", IEEE Mobile Multimedia Conference, 1999.
277. Matheson R, "The Electrospace Model as a Tool for Spectrum Management", http://www.its.blrdoc.gov/meetings/art/art03/slides03/mat_b/mat_abs.pdf, 2003.
278. Werback K, "Radio Revolution: The coming age of unlicensed wireless", www.werbach.com/docs/RadioRevolution.pdf, 2004.
279. Reed J, "Critical Choices for Designing Software Radios – SDR Forum DVD", 2002.
280. Notor J, "Cognitive Radio Emerges from Obscurity", Revision 2, <http://bwrc.eecs.berkeley.edu/Seminars/Notor%20-%201.23.04/Cognitive%20Radio%20Emerges%20from%20Obscurity%20Rev.2.pdf>, 2004.
281. Ling J, Trabelsi C, Valenzuela R, "Capacity of DECT Wireless Local Loop System in Multi-Cell Environment", 48th Annual Veh Tech Conf, 1998.
282. BTPLC, "World's First Converged Service" <http://www.btplc.com/Innovation/Mobility/converged>, 2004.
283. Neel J, Reed J, "Wireless, Wireless Everywhere and Now it Starts To Think", <http://www.eetimes.com/consumer/showArticle.jhtml?articleID=51200473>, 2004.
284. Sahai A, "Fundamental Limits on Cognitive Radio", www.eecs.berkeley.edu/wireless/posters/WFW05_cognitive.pdf, 2004.
285. Ofcom, www.ofcom.org.uk, [accessed 22 April 2005].
286. IEEE, "Improving Spectrum Usage Through Cognitive Radio Technology", <http://www.ieeeusa.org/policy/positions/cognitiveradio.asp>, 2003.
287. UniS/Guildford/UK, "Evaluation of Software Defined Radio Technology", SUR/CCSR/IND/OFCOM/SDRreport_v10, *to be published*.
288. PicoChip Designs Ltd, "picoChip unveils baseband technology to solve 3G infrastructure challenges", December 2002, http://www.picochip.com/press/press_releases/press005, [accessed 19 November 2004].
289. InfoWorld, "FCC approves first software-defined radio", 19 November 2004 http://www.infoworld.com/article/04/11/19/Hnfccradio_1.html, [accessed 19 November 2004].

UNCLASSIFIED

290. Michael L B, Mihaljevic M J, Haruyama S and Kohno R, "Security Issues for Software Defined Radio: Design of a Secure Download System", IEICE Trans Comms, vol E85 B, no 12, pp 2588, 2002.
291. Brawerman A, Blough D and Bing B, "Securing the Download of Radio Configuration Files for Software Defined Radio Devices", in Proc ACM 2nd International Workshop on Mobility Management and Wireless Access, Philadelphia, USA, October 2004.
292. SDR Forum, "Requirements for Radio Software Download for RF Reconfiguration", document number SDRF-02-A-0007-V0.0, 13 November 2002.
293. FCC, "Authorization and Use of Software Defined Radios – ET Docket no 00-47", FCC 01-264, 14 September 2001.
294. FCC, "Changes in certificated equipment", 47 CFR §2.1043, 1 October 2004.
295. FCC, "Identification of equipment", 47 CFR §2.925, 1 October 2004.
296. Directive 1999/5/EC of the European Parliament and of the Council of 9 March 1999 on Radio Equipment and Telecommunications Terminal Equipment and the Mutual Recognition of their Conformity, Official Journal L 091, pp 0010–0028, 7 April 1999.
297. Ofcom, European Commission Guidance Notes R&TTE Directive, February 2000, [downloaded from http://www.ofcom.org.uk/static/archive/ra/topics/conformity/document/rtte_rtteweb.htm, 30 November 2004].
298. Steinheider J, "Field Trials of an All-Software GSM Base Station", RF Design Mag, March 2004.
299. Ofcom, Sitefinder website, (<http://www.sitefinder.radio.gov.uk>, [accessed 22 April 2005]).
300. Ofcom, The Communications Market – Telecommunications Appendices, available at www.ofcom.org.uk, January 2005.
301. The Economist "Nokia's turnaround – The giant in the palm of your hand", 10 February 2005.
302. Nokia, "Texas Instruments Delivers Industry's First Integrated Single-Chip Solution for Mobile Phones; Nokia to Develop Mobile Phones Based on the Solution", Press Release, 24 January 2005 [downloaded from http://press.nokia.com/PR/200501/977043_5.html, 1 March 2005].
303. Texas Instruments, "Wireless Terminals Solutions Guide", 2Q2004 [downloaded from http://focus.ti.com/pdfs/wtbu/ti_wireless_solutions_guide.pdf, 22 April 2005].
304. Xilinx, "W CDMA RACH Preamble Detection Reference Design", Doc No: PN 0010822-1, 2005 [downloaded from www.xilinx.com, 1 March 2005].
305. Xilinx, "Xilinx Searcher Reference Design for W CDMA Node-B Receiver", Doc No: PN 0010857, [downloaded from www.xilinx.com, 1 March 2005].
306. Xilinx, "HSDPA Reference Design" [downloaded from <http://www.xilinx.com/esp/wireless/baseband/hsdpa.htm>, 1 March 2005].
307. Analog Devices, "TigerSHARC Embedded Processor – ADSP-TS201S", Datasheet, Rev. O, November 2004 [downloaded from www.analog.com, 2 March 2005].

UNCLASSIFIED

308. Rohde & Schwarz, "Universal Radio Communication Tester R&S® CMU200", Product Brochure, Version 6.00, May 2004 [downloaded from www.rohde-schwarz.com, 2 March 2005].
309. Texas Instruments Inc DSP Third Party Program, <http://dspvillage.ti.com/docs/thirdparty/thirdpartyhomepage.jhtml?templateId=5681>, [accessed 22 April 2005].
310. Xilinx Inc, www.xilinx.com, [accessed 22 April 2005].
311. Wong B, "Filling the Generation Gap with Software-Defined, Broadband Radio", CTI Magazine, vol 4, no 9, September 1999.
312. Tuttlebee W (ed), *Software Defined Radio – Origins, Drivers and International Perspectives*, John Wiley & Sons, Chichester, 2002.
313. Tan Z., "Comparison of Wireless Standards Setting – United States Versus Europe", <http://arxiv.org/ftp/cs/papers/0109/0109100.pdf>, January 2001.
314. CDMA Development Group website, www.cdg.org, May 2005.
315. GSM Association, "Growth of the Global Digital Mobile Market", www.gsmworld.com/news/statistics/pdf/gsma_stats_q4_04.pdf, May 2005.
316. HP Vanu, "Mid-Tex Cellular A Success Story from HP", www.vanu.com/docs/HP_MidTex_success.pdf, November 2003.
317. SDR Forum's Marketing Group, "SDR's Role in the Wireless World", www.govcomm.harris.com/dmtk/NeedForSDR.pdf, January 2000.
318. Bischoff Glenn, "Software Defined Radio Gets A Boost", http://mrtmag.com/mag/radio_technologies, December 2004.
319. Wickham Rhonda, "Long Awaited SDR Becomes Commercial Reality", Wireless Week, http://www.vanu.com/news/news/CTIA_2005.pdf, 16th March 2005.
320. Mobile Guru website, www.mobileguru.co.uk/Mobile_Technology_globe.html, May 2005.
321. AirNet website, www.aircom.com, May 2005.
322. Vanu website, www.vanu.com, May 2005.
323. Vanu, "Vanu, Inc. Announces Commercial Availability of Anywave Base Station", www.vanu.com/news/prs/anywave031405.pdf, 14th March 2005.
324. picoChip, "picoChip Technology Powers AS.MAX, Airspan's Family of WiMAX Base Stations Launched at CeBIT", www.picochip.com/press/press_releases/press031, March 2005.
325. picoChip site, www.picochip.com, May 2005.
326. Unstrung, "Airspan Scores on WiMax", www.unstrung.com/document.asp?doc_id=70025, 11th March 2005.
327. Haley Colin, "AT&T Tests WiMax Gear", DevXNews, www.devxnews.com/article.php/3492026, 22nd March 2005.
328. 3GNewsroom, "SpectruCell – SDR Base Station Presents Solution to US 3G Spectrum Rollout and Allocation Woes", www.3gnewsroom.com, May 2001.
329. 3GNewsroom, "ACT talking with the European to provide Shared/Virtual Networks Base Stations", www.3gnewsroom.com, 30th July 2001.

UNCLASSIFIED

330. Ofcom, "The Communications Market 2004 - Telecommunications", www.ofcom.org.uk/research/cm/cmpdf/telecoms.pdf, August 2004.
331. Cotton David, "Software Defined Radio Isn't Just About Software", COTS Journal, www.cotsjournalonline.com, January 2005.
332. Analogue Devices Inc., Press Release: "Analogue Devices' Wireless Technology Featured in GSM Association Award Winning Product", www.vanu.com/news/news/3GSM_2005_Award_Vanu_FINAL.pdf, Feb. 2005.
333. McCann Andy and Brannon Brad, "DSP Brings Base Station SDR Reality", RF design, www.rfdesign.com, September 2004.
334. PA Consulting, Press Release: "PA Consulting Group Develops the World's First all Software base band for W-CDMA 3G Base stations", www.paconsulting.com, September 2002.
335. 3G.co.uk, "TIs new 1GHz DSP Chip Supports Multiple 3G Standards", www.3g.co.uk, 11th February 2005.
336. Lucent Technologies, "A Guide to GSM Network Migration", www.lucent.com, April 2001.
337. Zhang et al, "Advanced Baseband Technology in Third Generation Radio Base Stations", Ericsson Review, www.ericsson.com, January 2003.
338. O'Shea Dan, "SDR Forum: Carriers Speak Out", Telephony Online, www.telephonyonline.com, 12th January 2004.
339. Walko John, "A Ticket to Ride", IEE Communications Engineer, Feb./Mar. 2005.
340. Mobile Pipeline, "Wi-Fi to hurt 3G Profits, Study says", Information Week, www.informationweek.com, 26th January 2005.
341. Kewney Guy, "WiMAX Roadshow Rolls into Town", The Register, www.theregister.co.uk, 25th October 2004.
342. WiMAX Forum website, www.wimaxforum.org, May 2005.
343. IEE Review, "Intel Consigns WiMAX to Science History", April 2005.
344. IEE Communications PN, "Airspan brings WiMAX closer to the Masses", www.iee.org/oncomms/sector/communications, 16th March 2005.
345. Wright Ian, "In Sync", IEE Communications Engineer, February/March 2005.
346. Sepura website, www.sepura.co.uk, May 2005.
347. Brown Alison, Gold Kenn and Nylund Mark, "A GPS Software Application for Embedding in Software Definable Radios", Proceedings of 13th Virginia Tech Symposium on Wireless Personal Communications, June 2003.
348. BBC News, "Car Charge Trials in '5 Years'", <http://news.bbc.co.uk>, 9th June 2005.
349. SAFECOM website, www.safecomprogram.gov, May 2005.
350. NTFI, "When They Can't Talk Lives Are Lost", www.agileprogram.org/ntfi/ntfi_brochure.pdf, February 2003.
351. Smith Brenna and Tolman Tom, "Can We Talk – Public Safety and the Interoperability Challenge", National Institute of Justice Journal, www.agileprogram.org/documents/jr000243d.pdf, April 2000.

UNCLASSIFIED

352. NPSTC website, www.npstc.org/whatisNPSTC.htm, May 2005.
353. Bischoff Glenn, "SDR Forum Targets Public Safety", MRT, www.mrtmag.com, 1st March 2004.
354. VDC, Press Release: "Public Safety Sector Calls for Software Defined Radio", www.vdc-corp.com/telecom/press/04/pr04-40.html, 2004.
355. Public Safety Wireless Network Programme, "Communications Lessons Learnt from the Pentagon Attack", www.safecomprogram.gov/NR/rdonlyres/8839D9BA-9104-4EE1-BC43-E8431C500F95/0/AnsweringCallLessonsPentagonAttack.pdf, January 2002.
356. PITO website, www.pito.org.uk, May 2005.
357. O2 Airwave website, www.airwaveservice.co.uk, May 2005.
358. O2 Airwave, Press Release: "Ofcom extends access to Airwave", www.airwaveservice.co.uk, 29th March 2005.
359. Firelink, "Firelink Bulletin", www.firelink.org.uk/resources/downloads/protected/newsletteroct.pdf, October, 2003.
360. AirNet, Press Release - "AirNet Receives \$1.4M Order for National Guard RapidCell Solution Improves Incident Site Command Capabilities", www.airnetcom.com/pressrel/natlguardorder.htm, 8th October 2004.
361. Vanu, SDR Product Concept - "Vanu Universal Public Safety Radio", [www.vanu.com/technology/Vanu Universal Public Safety Radio.pdf](http://www.vanu.com/technology/Vanu%20Universal%20Public%20Safety%20Radio.pdf), May 2005.
362. Waldschmidt C., Kuhnert C., Pauli M. and Wiesbeck W., "Handy MIMO", IEE Communications Engineer, February/March 2005.
363. QinetiQ, "JASMINE – Wideband Multi Frequency MIMO Channel Sounder", www.cpar.qinetiq.com/jasmine.html, May 2005.
364. Webb W., "Ofcom Opens the Door of Spectrum Management to the Forces of the Market", IEE Communications Engineer, February/March 2005.
365. Hart Chad, "Software Defined Radio (SDR): North America and European Market Demand Analysis", Venture Development Corporation website, www.vdc-corp.com, August 2004.
366. DECT Forum, "DECT – The Standard Explained", February 1997 [downloaded from <http://www.dect.org/pdf/TechnicalDocument.PDF>, 9 June 2005].
367. ETSI, Radio Equipment and Systems (RES); Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 2: Physical layer (PHL), ETS 300 175-2, September 1996.
368. Vodafone UK, "Hotspot List – 9th May 2005", 9 May 2005 [downloaded from [http://www.vodafone.co.uk/download/Vodafone%20WLAN%20\(BT%20Openzone\)%20hotspots_09May05.pdf](http://www.vodafone.co.uk/download/Vodafone%20WLAN%20(BT%20Openzone)%20hotspots_09May05.pdf), 14 June 2005].
369. T-Mobile, "T-Mobile HotSpot – High speed wireless internet access", http://www.t-mobile.co.uk/Dispatcher?menuid=phones_wb [accessed 4 July 2005].
370. BBC, "BT to launch fixed-mobile service", 15 June, 2005 [downloaded from <http://news.bbc.co.uk/1/hi/business/4094424.stm>, 16 June 2005].

UNCLASSIFIED

371. BBC (Wakefield J), "Cheaper mobile calls on the way?", 16 June, 2005 [downloaded from <http://news.bbc.co.uk/1/hi/technology/4096404.stm>, 16 June 2005].
372. BT Group, "BT Fusion", <http://www.btfusion.bt.com/index.aspx>, [accessed 16 June 2005].
373. Steele R (Ed), *Mobile Radio Communications*, John Wiley & Sons, Chichester, 1996.
374. Leaves P, Ghaheri-Niri S, Tafazolli R, Huschke J, Salter J and Ek M, "Performance Evaluation of Dynamic Spectrum Allocation for Multi-Radio Environments", Mobile Summit 2001, Barcelona, 10 12 September 2001 [downloaded from <http://www.ist-drive.org/papers/unis-mobilesummit2001-dsa-paper.pdf>, 20 June 2005].
375. Haartsen J C, "The Bluetooth Radio System", IEEE Personal Communications, vol 7, no 1, pp 28 36, February 2000.
376. Federal Communications Commision, "Notice of Proposed Rule Making and Order – In the Matter of Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies (ET Docket No. 03-108) and Authorization and Use of Software Defined Radios (ET Docket No. 00-47 (Terminated))", FCC 03-322, 30 December 2003.
377. Betz J W, "Feature Detection", 12 February 2003 [downloaded from <http://www.fcc.gov/realaudio/presentations/2003/021203/featuredetection.pdf>, 30 June 2005].
378. Hanzo L, Wong C H, Yee M S, *Adaptive Wireless Transceivers – Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA and OFDM Systems*, John Wiley & Sons Ltd, Chichester, 2002.
379. Carlson A B, *Communication Systems – An Introduction to Signals and Noise in Electrical Communications*, (3rd edition), McGraw-Hill, New York, 1986.

Abbreviations

1G	First Generation Mobile Networks
2G	Second Generation Mobile Networks
3G	Third Generation Mobile Networks
4G	Fourth Generation Mobile Networks
AAF	Anti-aliasing Filter
ABC	Always Best Connected
ACM	Adaptive Computing Machine
AD	Analog Devices
ADC	Analogue-to-Digital Converter
ADS	Advanced Design System
AGC	Adaptive Gain Control/Automatic Gain Control
ALU	Arithmetic Logic Units
AM	Amplitude Modulation
AMPS	American Mobile Phone System
AoA	Angle of Arrival
API	Application Programming Interface
ASCP	Application Specific Coprocessors
ASIC	Application Specific Integrated Circuit
ASIP	Algorithm Specific Instruction set Processor
ASK	Amplitude-Shift Keying
AT&T	American Telephone & Telegraph
AXU	Application specific eXecution Units
BB	Base-Band
BER	Bit Error Rate
BIST	Built-In Self Test
BPSK	Binary Phase-Shift Keying
BSC	Base Station Controller
BSS	Base Station Sub-System
BTS	Base Transceiver Station
C450	A German 1G standard
CA	Certification Authority
CDMA	Code Division Multiple Access
CF	Crest Factor

UNCLASSIFIED

CHAP	Challenge Handshake Authentication Protocol
CI	Carrier Interferometry
CINR	Carrier-to-Interference-plus-noise Ratio
CLB	Configurable Logic Block
CORBA	Common Object Request Broker Architecture
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
CR	Cognitive Radio
CRC	Communication Research Centre, Canada / Cyclic Redundancy Check
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DAB	Digital Audio Broadcast
DAC	Digital-to-Analogue Converter
DBPSK	Differential BPSK
DC	Direct Current
DECT	Digital Enhanced Cordless Telecommunications
DNL	Differential Nonlinearity
DoA	Direction of Arrival
DoD	Direction of Departure / Department of Defence
DOS	Denial Of Service
DPSK	Differential PSK
DRDC	Defence Research and Development Canada
DRiVE	Dynamic Radio for ip-service in Vehicular Environments
DSA	Dynamic Spectrum Allocation
DS-CDMA	Direct Sequence-CDMA
DSP	Digital Signal Processor
DSSS	Direct Sequence Spread Spectrum
DSTTD-SGRC	Double Space Time Transmit Diversity with SubGroup Rate Control
DTX	Discontinuous Transmission
D-TxAA	Double Transmit Antenna Array
DVB-H	Digital Video Broadcast-Handheld
DVB-T	Digital Video Broadcast-TV
EDGE	Enhanced Data Rates for Global Evolution
EJB	Enterprise Java Beans
EMC	ElectroMagnetic Compatibility
ENA	Electricity Networks Association

UNCLASSIFIED

ENOB	Effective Number Of Bits
ESSS	Earth Exploration Satellite System
ETSI	European Telecom Standardisation Institute
EU	European Union
FCC	Federal Communications Commission
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEA	Fractal Element Antenna
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FH-CDMA	Frequency Hopping CDMA
FHSS	Frequency Hopping Spread Spectrum
FIR	Finite Impulse Response
FPFA	Field-Programmable Functional Array
FPGA	Field Programmable Gate Array
FSA	Fixed Spectrum Allocation
FSK	Frequency Shift Keying
GFSK	Gaussian FSK
GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile communications
GUI	Graphical User Interface
HFSS	High Frequency Structure Simulator
HMI	Human-Machine Interface
HP	Hewlett Packard
HPA	High Power Amplifier
HR	Hardware Radio
HSDPA	High Speed Downlink Packet Access
HTTP	Hyper-Text Transfer Protocol
IDL	Interface Definition Language
IEEE	Institute of Electrical and Electronic Engineers
IFFT	Inverse FFT
IMD	Intermodulation Distortion
INL	Integral Non-Linearity

UNCLASSIFIED

IP	Intellectual Property / Internet Protocol
IPSec	Internet Protocol Security
IS-95	cdmaOne – US equivalent of GPRS
ISM	Industrial, Scientific and Medical
ISR	Ideal Software Radio
IST	Information Society Technologies
ITS	Intelligent Transport System
JTRS	Joint Tactical Radio System
JVM	Java Virtual Machine
LCD	Liquid Crystal Display
LED	Light-Emitting Diode
LMS	Least Mean Square
LNA	Low-Noise Amplifier
LO	Local Oscillator
LSB	Least-Significant Bit
MAC Ltd	Multiple Access Communications Limited
MAI	Multiple Access Interference
MC	Multi-Carrier
Mcps	Million Chips Per Second
MD5	Message Digest Algorithm #5
MDA	Model Driven Architecture
MEMS	Micro-Electromechanical System
MIMO	Multiple Input Multiple Output
MIPS	Million Instructions Per Second
MIT	Massachusetts Institute of Technology
ML	Maximum Likelihood
MLSE	Maximum Likelihood Sequence
MMAC	Multimedia Mobile Access Communication
MMI	Man-Machine Interface
MMSE	Minimum Mean Square Error
MNP	Mobile Number Portability
MOM	Message Orientated Model
MoU	Memorandum of Understanding
MRC	Maximum Ratio Combiner
MSps	Mega Samples Per Second

UNCLASSIFIED

MSRC	Modular Software-programmable Radio Consortium
MUD	Multi-User Detection
NFC	Near Field Communications
NMT	Nordic Mobile Telephone
NPRM	Notice of Proposed Rulemaking
NPSTC	National Public Safety Telecommunications Council
NSS	Network Sub System
NTIA	National Telecommunications and Information Administration
OBSAI	Open Base Station Architecture Initiative
OEM	Original Equipment Manufacturer
Ofcom	The Office of Communications
OFDM	Orthogonal Frequency-Division Multiplexing
OMA	Object Management Architecture
OMG	Object Management Group
OMRS	Operators Market Requirements Survey
O-QPSK	Offset QPSK
OSI	Open Standards Interface
OSIC	Ordered Successive Interference Cancellation
OSS	Operations Sub System
OVSF	Orthogonal Variable Spreading Factor
OWL	Web Ontology Language
P2P	Peer-to-Peer
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
PARC	Per-Antenna Rate Control
PC	Personal Computer
PDA	Personal Digital Assistant
PE	Processing Element
PHY	Physical Layer
PIM	Platform-Independent Model
PITO	Police Information Technology Organisation
PKI	Public Key Infrastructure
PM	Policy Model
PMR	Private Mobile Radio
PN	Pseudo Number

UNCLASSIFIED

PSD	Power Spectral Density
PSE	Problem Solving Environment
PSK	Phase-Shift Keying
PSM	Platform-Specific Model
PSMR	Public Safety Mobile Radio
PSTN	Public Switched Telephone Network
PU ² RC	Per-User Unitary Rate Control
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
R&D	Research & Development
R&TTE	Radio Equipment and Telecommunications Terminal Equipment
RAN	Radio Access Network
RC MPD	Rate-Control Multi-Path Diversity
RDF	Resource Description Framework
RDS	Radio Data Service
RF	Radio Frequency
RLS	Recursive Least Squares
RMS	Root Mean Square
ROC	Recovery Orientated Computing
ROM	Resource Orientated Model
RSFQ	Rapid Single Flux Quantum
RSM	Radio Security Module
RSSI	Received Signal Strength Indication
RTMS	An Italian 1G standard
RTOS	Real-Time Operating System
SBC DTF	Software Based Communication Domain Task Force
SCA	Software Communications Architecture
SCARISCA	reference implementation
SCR	Software Controlled Radio
SDR	Software Defined Radio
SDRF	Software Defined Radio Forum
SFBC	Space-Frequency Block Coding
SFDR	Spurious-Free Dynamic Range
SHA1	Secure Hash Algorithm 1

UNCLASSIFIED

SINR	Signal-to-Interference-plus-Noise-Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input, Single-Output
SMS	Simple Messaging Service
SNR	Signal-to-Noise Ratio
SOA	Service Orientated Architecture
SOAP	Simple Object Access Protocol
SoC	System-on-Chip
SOM	Service Orientated Model
S-PARC	Selective Per-Antenna Rate Control
Sps	Sample per second
SQNR	Signal-to-Quantisation Noise Ratio
SRR	Software Reconfigurable Radio
SSL	Secure Sockets Layer
SSPA	Solid-State Plasma Antenna
STBC	Space-Time Block Codes
STTC	Space-Time Trellis Codes
STTD	Space-Time Transmit Diversity
ST-TuC-turbo-PIC	Space-Time Turbo-Coded Turbo Parallel Interference Cancellation
TACS	Total Access Cellular System
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TETRA	TERrestrial Trunked Radio
TI	Texas Instruments
TPC	Transmit Power Control
TRAU	Transcoder Rate Adaptor Unit
TSTD	Time Switched Transmit Diversity
UDDI	Universal Description and Discovery Integration
UK	United Kingdom
ULA	Uniform Linear Antenna array
UML	Unified Modelling Language
UMTS	Universal Mobile Telecommunications System
URI	Universal Resource Identifier
US	United States
USR	Ultimate Software Radio

UNCLASSIFIED

UTRA	UMTS Terrestrial Radio Access
UWB	Ultra Wide Band
VCO	Voltage-Controlled Oscillator
VDC	Venture Development Corporation
VHDL	Very High Speed Integrated Circuit Description Language
VM	Virtual Machine
VO	Virtual Organisation
VoIP	Voice-over-IP
VPN	Virtual Private Network
VSWR	Voltage Standing Wave Ratio
W3C	World Wide Web Consortium
WCDMA	Wideband Code Division Multiple Access
WDL	Waveform Description Language
WHT	Walsh-Hadamard Transform
WiFi	IEEE 802.11 wireless local area network compliant products
WiMax	IEEE 802.16 wireless metropolitan area network compliant products
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WPAN	Wireless Personal Area Network
WSA	Web Services Architecture
XG	neXt Generation
XML	eXtensible Mark-up Language
ZF	Zero-Forcing

A Abstracts from SDR and MIMO Search

Terms: ("SDR" OR "Software Radio") AND "MIMO"

Author(s): Becher, R., Dillinger, M., Haardt, M., Mohr, W.

Title: Broad-band wireless access and future communication networks

Source: PROCEEDINGS OF THE IEEE, vol. 89, no. 1, pp. 58-75, January 2001

Abstract: This paper presents a vision for wireless communication systems beyond the third generation, which comprises a combination of several optimized access systems on a common IF-based medium- access and core network platform. These different access systems will interwork via horizontal and vertical handover, service negotiation, and global roaming. The different access systems are allocated to different cell layers in the sense of hierarchical cells with respect to cell size, coverage, and mobility to provide globally optimized seamless services to all users. This vision requires extensive international research and standardization activities to solve many, technical challenges. Key issues are the global interworking of different access systems on a common platform, the implementation of multimode and multiband terminals and base stations by software-defined radio concepts as well as advanced antenna concepts.

Author(s): Schnepp, H., Muller, T., Luy, J., F., Russer, P.

Title: The implementation of channel diversity in mobile software radio receivers

Source: IEEE MICROWAVE AND WIRELESS COMPONENTS LETTERS, vol. 13, no. 8, pp. 323-325, August 2003

Abstract: A novel channel diversity concept is proposed and demonstrated, which avoids receiving signal deterioration due to multipath fading in mobile receivers. The system is based on coherent superposition of the signals received from several transmitters supplying the same information at different frequencies. Based on a software radio architecture this concept may increase the quality of mobile reception in modern car receivers considerably. Compared with multiantenna-receivers which overcome the multipath fading problem by simultaneously receiving the same program with several antennas, the proposed solution is advantageous, since it requires only a single antenna.

Author(s): Gifford, S., Kleider, J., E., Chuprun, S.

Title: OFDM-MIMO communication systems in a Rayleigh faded environment with imperfect channel estimates

Source: IEEE MILCOM 2003, 2003 IEEE Military Communications Conference, IEEE cat. no. 03CH37500, vol. 1, pp. 633-637, 2003

UNCLASSIFIED

Publisher: IEEE, Piscataway, NJ, USA

Abstract: This paper presents the performance of mobile orthogonal frequency division multiplexing (OFDM) multiple-input multiple-output (MIMO) communication systems with imperfect knowledge of the channel matrix. MIMO systems typically require a channel matrix, which can be determined initially from a training sequence. However, mobile communication systems exhibit a time-varying channel matrix and have time and frequency selective fades which result in performance degradation of the MIMO system. Channel tracking methods can be used to estimate the time-varying channel matrix but cannot in practice be error free. This paper presents results of V-BLAST (vertical Bell Laboratories layered space-time) MIMO simulations using the geometric wideband time-varying channel model (GWTCM) with Rayleigh faded environments and imperfect channel matrix knowledge. Flat fading is assumed for each OFDM subcarrier. OFDM-MIMO architectures such as OFDM coupled with V-BLAST can be easily implemented by exploiting the built-in and flexible multi-channel architectures of advanced software defined radios (SDR).

Author(s): Morawski, R.. Le-Ngoc, T., Naeem, O.

Title: Wireless and wireline MIMO testbed

Editor(s): Oliver, G., Pierre, S., Sood, V., K.

Source: IEEE CCECE 2003, Canadian Conference on Electrical and Computer Engineering, Toward a Caring and Humane Technology, IEEE cat. no. 03CH37436, vol. 3, pp. 1913-1916, 2003

Abstract: This paper presents the multiple-input-multiple-output (MIMO) testbed for rapid characterization of practical MIMO wireless and wireline channels, as well as verification, performance evaluation and demonstration of newly developed transmission and signal processing algorithms in real life environments. The testbed is ideal for development of advanced software-defined radio (SDR), wireless and wireline SISO, SIMO, MSO and MIMO systems. Detailed architecture, specifications and operation of the MIMO wireless and wireline testbed, and some experimental results are discussed.

Author(s): Eireiner, T., Muller, T., Luy, J., F., Owens, F.

Title: Implementation of a smart antenna system with an improved NCMA algorithm

Source: 2003 IEEE MTT S International Microwave Symposium Digest, IEEE cat. no. 03CH37411, vol. 3, pp. 1529-1532, 2003

UNCLASSIFIED

Abstract: This paper presents the possibility of using adaptive algorithms for digital beamforming purposes. A normalized constant modulus algorithm (NCMA) is implemented in a standard FPGA. In this way, a simple and non-hardware-intensive, smart antenna system in combination with a software-defined radio (SDR) has been realized for mobile FM reception. A new kind of algorithm initialization, leads to an improvement in startup behaviour. The quality in signal separation makes the NCMA algorithm also suitable for MIMO purposes. The NCMA algorithm increases the reception quality for mobile communication systems dramatically.

Author(s): Rao, R., M., Daneshrad, B.

Title: I/Q mismatch cancellation for MIMO-OFDM systems

Source: 2004 IEEE 15th International Symposium on Personal Indoor and Mobile Radio Communications, IEEE cat. no. 04TH8754, vol. 4, pp. 2710-2714, 2004

Abstract: MIMO-OFDM systems are gaining prominence in high data rate applications due to their increased spectral efficiency, in MIMO-OFDM systems, I/Q mismatch causes interference between the frequency mirror subcarriers. This can significantly degrade and limit the performance of the system. In this paper, we present a mathematical analysis of the effects of I/Q gain, delay and phase mismatches in MIMO-OFDM systems. We show the impact of I/Q mismatch when more antennas are added. We also derive a linear optimal solution to cancel all three I/Q mismatches and an RLS based adaptive filter for the same. We show simulation results and results on an experimental testbed.

Author(s): Stüber, G., L., Barry, J., R., McLaughlin, S., W., Li, Y., Ingram, M., A., Pratt, T., G.

Title: Broadband MIMO-OFDM wireless communications

Source: PROCEEDINGS OF THE IEEE, vol. 92, no. 2, pp. 271-294, February 2004

Abstract: Orthogonal frequency division multiplexing (OFDM) is a popular method for high data rate wireless transmission. OFDM may be combined with antenna arrays at the transmitter and receiver to increase the diversity gain and/or to enhance the system capacity on time-varying and frequency-selective channels, resulting in a multiple-input multiple-output (MIMO) configuration. The paper explores various physical layer research challenges in MIMO-OFDM system design, including physical channel measurements and modelling, analogue beam forming techniques using adaptive antenna arrays, space-time techniques for MIMO-OFDM, error control coding techniques, OFDM preamble and packet design, and signal processing algorithms used to perform time and frequency synchronization, channel estimation, and channel tracking in MIMO-OFDM systems. Finally, the paper considers a software radio implementation of MIMO-OFDM.

Author(s): Araki, K.

UNCLASSIFIED

Title: RF analogue smart devices/circuits and their applications

Source: Transactions of the Institute of Electronics, Information and Communication Engineers, vol. J87-C, no. 1, pp. 3-11, January 2004

Abstract: RF analogue controllable devices are strongly required in order to realize advanced mobile communication and radar/sensing systems, where a sophisticated space-time signal processing will be widely employed. These devices are also necessary for high precision measurement systems. We address these principles, practical examples and forthcoming application systems with RF analogue controllable devices in this paper.

B Selected Google Search Results for SDR and MIMO

Terms: “SDR” AND “MIMO”

1. Rapid MIMO-OFDM Software Defined Radio System Prototyping

This paper describes a MIMO-OFDM prototype built using a flexible, software defined radio (SDR) system architecture in conjunction with commercially available ...

<http://www.ece.utexas.edu/~rheath/papers/2004/sips/>

2. Rapid MIMO-OFDM Software Defined Radio System Prototyping

The flexibility enabled by an SDR MIMO- OFDM prototype becomes clear in the following section where we present a detailed description of the prototyping platform ...

<http://www.ece.utexas.edu/~rheath/papers/2004/sips/paper.pdf>

3. N04-109 - Universal MIMO-OFDM SDR for Mobile Autonomous Networks

Universal MIMO-OFDM SDR for Mobile Autonomous Networks Navy SBIR FY2004.1. ... Topic Title: Universal MIMO-OFDM SDR for Mobile Autonomous Networks. ...

http://www.navysbir.com/04_1/213.htm

4. III. Project Description

... digital components, leading to the so-called “software defined radio” (SDR) technology. ... the development of the multi-input multi-output (MIMO) technology. ...

<http://www.eic.nctu.edu.tw/ACE/a3/a392content.pdf>

5. Microelectronics and Information Systems research Centre

... and Testing. B3G Wireless Access Technology: SDR/OFDM/MIMO. Network Benchmarking Laboratory: Performance/Security/Inter-Operability.

<http://www.eic.nctu.edu.tw/ENG/itri/itri.htm>

6. SDR '04 - Exhibitor Information

... They provide unique capabilities for testing SDR systems, MIMO (up to 4x4) and smart antenna products in terrestrial or airborne environments up to missile ...

http://www.SDRforum.org/SDR04/exhibit_notice.html

7. Prototyping advanced 3G/4G wireless and SDR (Software Defined ...

... Chuprun, General Dynamics, Decision Systems, “A flexible geometric wide-band time-varying channel model for V-Blast MIMO simulations”, SDR’02 Conference ...

http://www.lyrtech.com/fr/publications/article_prototyping_advanced_3g_4g_en.pdf

UNCLASSIFIED

8. Keynote: Software Defined Radio platform
Tutorial SympoTIC'04 An Information Theoretic Point of View to MIMO Channel Modelling Merouane Debbah Institute Eurecom, Mobile Communications Group 2229 ...
<http://www.ctl.elf.stuba.sk/sympotic04/tutorial.html>
9. CRC Strategic Research Plan 2004-2007
... Software Defined Radio (SDR), MIMO and innovative network topologies. Test beds are developed/maintained to perform engineering testing when required.
...
http://www.crc.ca/en/html/crc/home/info_crc/strategic_research_plan/radio_spectrum_e.html
10. Hot issues on UWB, MIMO and Security for Wireless Broadband ...
Hot issues on UWB, MIMO and Security for Wireless Broadband Technologies (Last Minute Session on Wireless Communications). Date: Thu ...
<http://www.ieice.org/cs/ap/ISAP2002/contents/lms1.html>
11. Microelectronics and Information Systems research Centre
... and Testing. B3G Wireless Access Technology: SDR/OFDM/MIMO. Network Benchmarking Laboratory: Performance/Security/Inter-Operability.
<http://www.eic.nctu.edu.tw/ENG/itri/itri.htm>
12. Smart Antenna Overview An Introductory Presentation for SDR Forum
... signal and 802.11 based WLAN physical layer signals • MIMO provides throughput enhancement; suitable for the upcoming ... Lends itself to SDR implementation ...
http://www.SDRforum.org/MTGS/mtg_39_jun04/04_i_0060_v0_00_api_factors_06_07_04.pdf
13. COBRA - COnputer Based Radio and Antennas
... space-time signal processing approaches for systems with multiple antennas at the transmitter and the receiver (Multiple Input Multiple Output channels - MIMO ...
http://www.ifn.et.tu-dresden.de/MNS/research/projects_cobra.html
14. Hot issues on UWB, MIMO and Security for Wireless Broadband ...
Hot issues on UWB, MIMO and Security for Wireless Broadband Technologies (Last Minute Session on Wireless Communications). Date: Thu ...
<http://www.ieice.org/cs/ap/ISAP2002/contents/lms1.html>
15. Wireless Breakthrough - By Marc Beacken, Alex Pidwerbetsky and ...
... The MIMO SDR is a scalable, evolvable, integrated, multiprocessor platform with extensive FPGA support for computationally demanding modem and radio ...
http://www.military-information-technology.com/archive_article.cfm?DocID=141
16. Prototyping a MIMO W CDMA system using a system-level approach

UNCLASSIFIED

- ... project at Laval university involving the efficient implementation of MIMO algorithms for ... logic parts and the growing importance of SDR / reconfigurable radios ...
- http://www.lyrtech.com/fr/publications/Prototyping_MIMO-WCDMA-system.pdf
17. The 6th Smart Antenna Workshop
... system). 11:10~11:30, Sandbridge SDR Processor Technology with Application to MIMO-OFDM Receiver, Dr. John Glossner (Sandbridge, USA). ...
<http://dsplab.hanyang.ac.kr/6th%20workshop%20for%20detail.html>
18. MIMO Wireless Communications Research
... MIMO OFDM Testbed Development. - Three phase approach that starts with a non real-time SDR based approach and culminates in the Gbps 8x8 system. ...
<http://www.MIMO.ucla.edu/>
19. ::::::::::::KOREA ETHERNET FORUM:::::::::::
... Session ?. 14:00 ~ 14:30 SDR Requirements - Dr. Mamoru Sawahashi (NTT DoCoMo, Japan). ... 15:00 ~ 15:30 Precoding techniques for MIMO channels - Prof. ...
<http://www.ethernet.or.kr/pg003-02.htm>
20. MWSCAS 2004
... GMPLS); HDR Modulation and Coding Techniques; Network Security; Next Generation Wireless Systems (SDR/MUD/MIMO/OFDM); Ultra-Wideband ...
<http://www.huis.hiroshima-u.ac.jp/mwscas04/topics.html>
21. Microsoft PowerPoint - Anaheim Slides - Seesta
... processing power could be used also for multi-band SDR and –MIMO – maybe even SoC will be possible? – Makes it possible to build a very robust system
http://www.ieee802.org/21/may04_meeting_docs/21-04-0043-00-0000-Seesta.pdf
22. ::::::::::::KOREA ETHERNET FORUM:::::::::::
... Session ?. 14:00 ~ 14:30 SDR Requirements - Dr. Mamoru Sawahashi (NTT DoCoMo, Japan). ... 15:00 ~ 15:30 Precoding techniques for MIMO channels - Prof. ...
<http://www.ethernet.or.kr/pg003-02.htm>
23. SiPS 2004
... is of great importance.This paper describes an approach for prototyping a MIMO-OFDM system using a flexible software defined radio (SDR) system architecture in ...
<http://www.spsworkshops.com/SiPS2004/Papers/viewpapers.asp?papernum=1089>
24. Polarizone Technologies Sdn Bhd
... innovative modulations , multiplexing, MIMO processing , space time coding and smart antennas . Scientist - MIMO group (Kuala Lumpur). ...

UNCLASSIFIED

- <http://my.jobstreet.com/jobs/2004/10/default/10/148522.htm>
25. 3Gwireless'01
... concepts 5. Standardisation and regulatory topics for SDR terminals T05 ... Channels, Fading, Diversity • Multiple-Input Multiple-Output (MIMO) Wireless Systems 2 ...
<http://delson.org/3gwireless01/tutorial.htm>
26. 17.12.2004
... Methodologies for Signal Processing Algorithms and the INTIHT at TU Wien gave an overview of possible applications of SDR with more than ... 12.2.2003 MIMO Meeting, ...
<http://www.nt.tuwien.ac.at/cdlab/news.htm>
27. SDR-3000 MRDP
... Waveform Development • Verification and Validation testing • Beamforming and MIMO research ... of up to 16 MHz • Based on Spectrum's SDR-3000 platform ...
http://www.spectrumsignal.com/products/pdf/SDR_3000_mrpd.pdf
28. HY-SDR Research Center
... SAMSUNG Electronics, Korea). Sandbridge SDR Processor Technology with Application to MIMO-OFDM... Dr. John Glossner (Sandbridge, USA).
<http://dsplab.hanyang.ac.kr/workshop/work6th/session1/sub.htm>
29. Broadband MIMO-OFDM Wireless Communications
Page 1. Broadband MIMO-OFDM Wireless Communications GORDON ... systems. Finally, the paper considers a software radio implementation of MIMO-OFDM. ...
<http://users.ece.gatech.edu/~barry/pubs/journal/MIMO-ofdm.pdf>
30. Belkin to ship first consumer MIMO products : TomsNetworking ...
... New Products: Belkin to ship first consumer MIMO products Posted by Tim [2004-08-09 09:12:18] Send this story to a friend Printer friendly page. ...
http://www.tomsnetworking.com/News_story_729.php

C Web Services Security Specifications

Issues relating to security and the user's confidence in it will be a significant set of problems that will need to be addressed in order to assure the long-term adoption of SDR. The Wide Web Consortium (W3C) is coordinating efforts towards a set of open and comprehensive standards to tackle the challenges that are faced when forming and managing service orientated architectures (SOAs), to allow the negotiation of security issues and the resolution of conflicts that arise in ad-hoc arrangements. These web services security specifications are summarised in the following sections.

A single sign-on authentication can be used once and passed through a chain of web services to complete composed services as if it were single transaction. In such cases, secure sockets layer (SSL)/transport layer security (TLS) and public key infrastructure (PKI) could be used to ensure transport level security. However, messages must be decrypted by each service and there is no method to verify that the received message is the same as that sent. The WS-Security and supported specifications aim to detail mechanisms by which services can establish, or assert, the trustworthiness of messages they receive and send.

Web service security is split into four layers, comprising seven specifications, with a foundation of SOAP messaging as shown in Figure C 1.

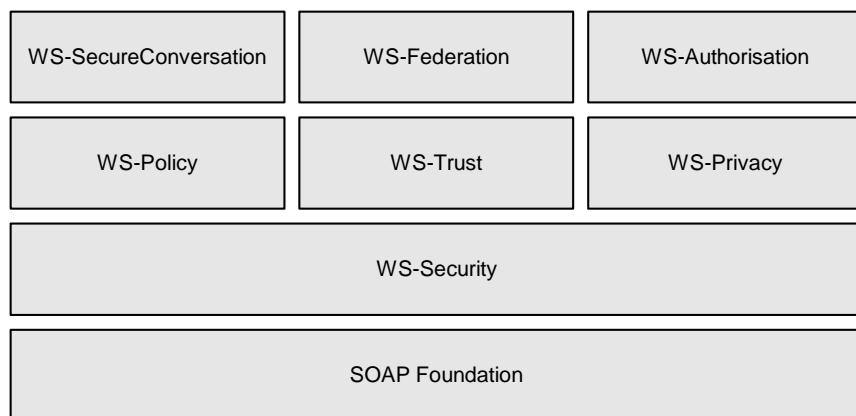


Figure C 1: Web services security specification.

These specifications allow for message level security in web service and grid service communication and can be used in a modular fashion, rather than a service being required to conform to all the specifications in order to offer a small piece of secure functionality.

WS-Security

The aim of the WS-Security specification is to provide applications, specifically web service applications, with a method of constructing secure SOAP message exchanges. To this end, WS-Security describes SOAP extensions to provide Quality of Protection for web service messages and methods to encode binary security tokens (specifically X509 certificates and Kerberos tickets) within SOAP headers.

UNCLASSIFIED

Quality of Protection comprises message integrity, which is achieved using the XML signature mechanism with security tokens and message confidentiality, which is achieved through the XML encryption mechanism and security tokens. The WS-Security specification also provides an extensible mechanism for associating any number of security tokens with a message that can be contained within the message itself or located at another endpoint (e.g., at another web service providing security tokens). The specification does not specify any requirements about the security token themselves, thus allowing for any implementation to be used.

WS-Security provides a common base functionality for all the other WS specifications and is aimed to provide Message Level Security. Considerations of Transport Level Security are not covered by WS-Security, but may be addressed in other documents such as WS-Policy.

WS-Policy

WS-Policy specifies how senders and receivers of messages can state their capabilities and any requirements that must be met to allow communication to take place. Such things include what (if any) security tokens are required by the service and which encryption algorithms are supported.

A policy can contain a number of alternatives which are different sets of assertions that the service will accept to allow communication. An example is that the service will accept either a Kerberos or X509 security token.

WS-Policy also allows for the intersection of two or more policies from different domains to produce a single policy that will allow communication with members of any of the included domains upon successfully fulfilling the requirements of the intersected policy. Such a function is required if ad-hoc SDR networks are to be formed and are to interact in a transparent and seamless way with other services.

WS-Policy does not specify anything about how a policy is discovered by, or attached to, a service or resource. This is left to the WS-PolicyAttachment specification. However, it does recommend that all policies and assertions are signed to prevent tampering and that no policy should be accepted unless it is signed and is provided with a security token indicating the signer has sufficient privileges to speak for the policy's scope.

WS-Trust

The WS-Trust specification is currently in an initial draft stage. However, it is intended to provide a framework to allow services and requesters to establish secure, trusted communications. It is built upon the WS-Security specification and defines extensions to allow for security token exchange (including issue, renewal and validation) and for the establishment and assessment of trust relationships (both direct and brokered) between applications, services or resources.

WS-Trust includes a security token service (STS) framework that defines RequestSecurityToken and RequestSecurityTokenResponse elements. Also included is a set of bindings for the elements that specify the permitted extensibility and composition with other parts of WS-Trust when issuing, renewing and validating security tokens.

UNCLASSIFIED

A STS, which has a policy detailing what it requires in order to provide a security token, would receive a RequestSecurityToken message from an entity that meets the service's requirements (as specified in its policy) and, if the policy permits, the STS would respond with a RequestSecurityTokenResponse that contains the required security token.

As well as this simple scenario, WS-Trust defines extensions to enable negotiation and challenges between a STS and a security token requester, using multiple iterations of sending RequestSecurityTokenResponse elements with either the challenge or reply contained in them. Once the challenges are completed to the satisfaction of the STS, the final RequestSecurityTokenResponse is sent with the token embedded.

In both of these scenarios, the requester could be an entity that wishes to use a particular service and which specifies in its policy that a security token from a specific STS is required. Alternatively, it could be the service itself that attempts to establish the trustworthiness of a requester with the STS. In both cases, the service explicitly trusts the STS, which is acting as a trusted intermediary between the two entities.

STSS form the basis of trust brokering. It is quite possible that several STSS could be queried in order to establish trust between two entities.

WS-Privacy

The WS-Privacy specification has not yet been written. However, it is intended to develop a privacy language that will allow services to specify privacy requirements as part of their policy. The privacy language will also allow services to make claims about their adherence to various privacy policies. By combining the use of WS-Policy to state a set of privacy requirements, WS-Trust to provide trustable claims of adherence to a privacy policy, and WS-Security to associate privacy claims with a message, an extensible and flexible privacy framework can be maintained.

WS-Privacy could be of great benefit within a federation of SDR networks, where varying levels of access to information localised to particular equipment might be required. Access could be according to an entity's role, location or any context readily modelled in a supporting ontology. An example would be details of secondary communication paths (such as engineering order wires) or fault tolerance strategies where it might not be acceptable for these data to be visible to the network as a whole. By utilising WS-Privacy statements within their policy, an individual could restrict access of the network as a whole to only being able to request higher level information. At the same time, their privacy settings could allow users from a shared sub-domain to view more sensitive and detailed information. Such access control would be specified in policy alternatives, each mapping to a set of access permissions within the user's SDR applications and configurations.

WS-SecureConversation

WS-SecureConversation is currently in revised public draft release and addresses a specific limitation of WS-Security, namely that it only provides single-message security. It is highly likely that two or more entities will wish to engage in a multi-message exchange without having to authenticate themselves to each other with each message that is passed. WS SecureConversation aims to create security contexts (SCs) and security context tokens (SCTs) to enable secure multi-message communication. SCTs contain a set of claims that are valid for the lifetime of the SC. They may be amended by any party trusted by the SC to include additional claims, but cannot have claims removed.

A SC needs to be created and distributed amongst the interested parties before it can be used. The WS-SecureConversation specification outlines three distinct mechanisms through which this can be achieved .

1. The SCT is created by a SCT service at the request of the conversation initiator. The initiator then distributes the SCT to the other parties in the conversation. The SCT service must be trusted by all participants in the conversation. This is shown in Figure C 2.

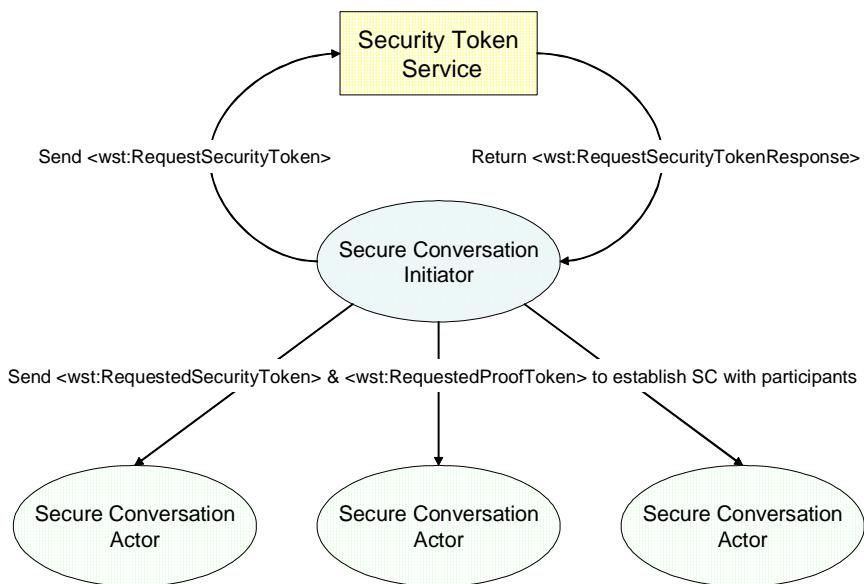


Figure C 2: SCT created by service

2. If the conversation initiator is trusted by all parties, they can create the SCT and distribute it with the messages sent to initiate the conversation. The other participants then decide whether or not to accept the SCT and enter into the secure conversation. This is shown in Figure C 3.

UNCLASSIFIED

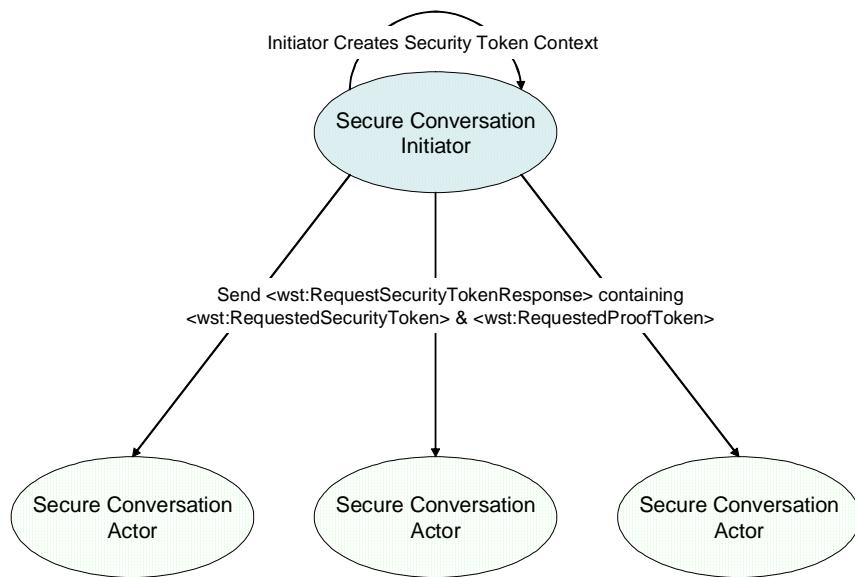


Figure C 3: Conversation initiator creates SCT

3. The SCT is created through negotiation between the various parties involved. An example might be the need to establish a shared secret between the participants. Any number of challenge/response messages can be sent following the framework described in WS-Trust to finally establish the SCT to be used between the conversation participants. This is shown in Figure C 4.

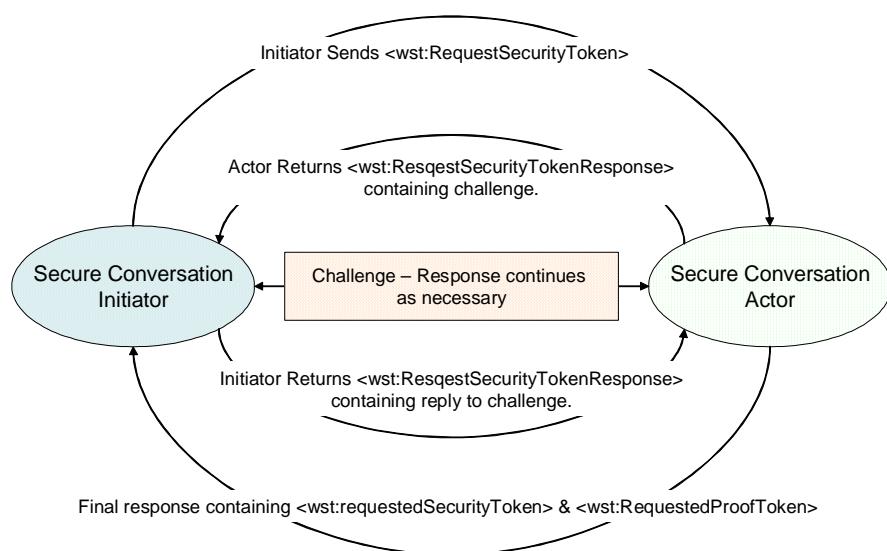


Figure C 4: SCT created through negotiation

UNCLASSIFIED

As with the other WS-* specifications, WS-SecureConversation provides an extensible framework and flexible syntax to enable implementation of a variety of security mechanisms and does not define any such mechanisms itself. To further secure a conversation, WS SecureConversation provides a mechanism for using derived keys once a SC has been established, which prevents the unnecessary resending of the original key information.

WS-Federation

The WS-Federation specification is currently in initial public draft release. It describes how to combine the use of WS-Trust, WS-Security and WS-Policy to create trust realm mechanisms for use between and within federations, thus providing for management and brokerage of trust in federated, heterogeneous environments.

Once within a federated environment, users/entities will often wish to keep their true identity and much of their personal information private from a number of different users or groups. To this end, WS-Federation introduces identity providers (IPs), attributes and pseudonyms and can be combined with WS-Privacy to restrict access to such information to trusted parties only.

IPs are very similar to STSs and might often be combined into the same service. An IP will, at a minimum, provide peer entity authentication and can provide identity claims about the authenticated peer in security tokens. IPs are expected to act as proxies on a requester's behalf to enable single sign-on. Such IPs must be trusted by the requester to maintain the secrecy of their identity information and by the resource to provide accurate, reliable identity information about the requester.

Two components of an IP are attribute and pseudonym services. Attribute services allow information about an entity to be stored and retrieved by anyone with sufficient, acceptable credentials, with different attributes able to have differing credential requirements for access. A pseudonym service is a special case of an attribute service that maintains and provides alternate identity information and, possibly, associated security tokens for entities within an attribute service. Pseudonyms and their associated tokens are scoped to specific services and an IP. This allows for more efficient authentication and allowing greater privacy options with a user only revealing their true identity to their trusted IP when they request the security token(s) rather than revealing it to each service they come across, possibly over insecure channels.

In some instances, a service might want to indicate where a requester could go to obtain the security tokens required to satisfy the claims requirements specified in its policy. WS Federation defines a RelatedService element that can be used to specify the endpoint reference of an IP, attribute service, pseudonym service or STS, that will meet its requirements.

UNCLASSIFIED

The final aspect of federated services addressed in WS-Federation is that of sign out. When users sign out from a service or federation, they indicate that it is now permissible to clear any cached information relating to them such as security tokens or state information. Sign out is intended to be provided by the IP/STS, as they act as a single point of contact for the user and all the services that they are currently utilising. In a federated environment, a sign out message will often need to be propagated throughout the federation to ensure that the user's cached data are fully removed from the federation. To enable this, WS-Federation specifies RequestSSOMessageEndpoint, RequestSSOMessage and CancelSSOMessage elements that are used by services to express an interest in receiving sign out messages relating to their realm and, optionally, for specified security tokens (i.e., only relating to a specified entity, or for all sign out messages for the services realm).

WS-Authorisation

The WS-Authorisation specification has not yet been released. However, it is expected to describe how access policies for services are specified and managed. More specifically, it is expected to state how claims should be made within security tokens and how they are to be processed at an endpoint.

Authorisation is concerned with what an actor is allowed to do once identity has been established. Currently, most systems have a simple allow/deny approach to remote authorisation. However, with the development of more complex web and grid service based environments, a more complex permission structure will be required. Some approaches, such as the CAS and the virtual organisation management system (VOMS), are already in use. However, there is currently no accepted standard as yet to ensure interoperability in line with the web service ideals.

Security Assertion Mark-up Language

The security assertion mark-up language (SAML) defines a language for the communication of identity information between business partners in a secure fashion. This mechanism is needed because browser and web service transactions blur the boundaries between separate business partners due to the flow of application data between them. Thus, identity management mechanisms must flow across these boundaries as well, accompanying the fundamental transaction data. Traditional authentication systems have required enterprises to maintain a one-to-one mapping of identity within their business systems for their customers, suppliers and partners. In this model of identity management, customer identity data must be registered and maintained within the enterprise's electronic authentication databases. This model, with this relatively tight coupling of identity data between business partners, does not easily scale to support today's web service and grid infrastructures. Thus, SAML provides a scalable mechanism for identity communication.