Future Mobile Telecommunications Systems

Peter F. Driessen
Department of Electrical and Computer Engineering
University of Victoria
Victoria, B.C., Canada
driessen@uvic.ca

ABSTRACT: In this paper, we review 2nd and 3rd generation wireless telephony systems, both in terms of technology and service concepts. We point out some of the technical and service issues that may affect the functionality and performance from the user point of view. In particular, we focus on one technical and one service concept important for long term success of 3rd generation systems: transmit diversity to increase capacity (quantity),and context-sensitive applications to increase user acceptance (quality).

1. Introduction

Wireless telephony and the Internet are both expanding rapidly, and finding new users around the world. Future mobile telecommunication systems will converge with the Internet and corporate Intranets to provide integrated voice/data/video (multimedia) 'anywhere, anytime' communications services to people on the move. The market demand for these services is huge, and presents both opportunities and challenges.

The challenge is two-fold: to provide genuinely useful information access and management services which are as easy to use as a telephone (i.e. quality), and to provide the system capacity to send enough bits per second per Hertz per unit area to accommodate the peak and average data rates needed to support these services (i.e. quantity). In this paper, we review two key ideas to meet these challenges: context-sensitive services for quality, and transmit antenna diversity for quantity.

Wireless telephone systems may be broadly divided into three generations. The first generation systems such as NMT (Nordic Mobile Telephone) and AMPS (Advanced Mobile Phone Service) use analog FM (Frequency Modulation) technology, provide only voice communications service, and were originally marketed to business users. These systems are still widely used today, and can accommodate limited data communications using the audio channel and a modified telephone-type modem.

The second generation systems such as IS-136 TDMA (Time-Division Multiple Access) with CDPD (Cellular Digital Packet Data), IS-95 CDMA (code-division multiple access) and GSM (Global System for Mobile)/PCS1900 (Personal Communications System) use digital modulation technology (i.e. voice is digitized using a vocoder and transmitted as data), provide short messaging service (SMS), low speed data, call management features in addition to voice service, and are marketed to consumers as well as business users. These second generation systems are currently being deployed, and have been in service for only the last few years. In some areas, the first and second generation systems compete with each other.

There are also 'low-tier' second generation systems with limited coverage intended for pedestrian-only (no vehicular) use. Such systems include PHS (Personal Handyphone System), CT-2 (Cordless Telephone 2), DECT (Digital European Cordless Telephone), and provide primarily voice service to consumers.

The third generation system IMT-2000 (International Mobile Telephone - 2000) is planned to be a single worldwide standard or group of standards while accommodating a variety of terminal types (voice/data/video), using more advanced digital modulation technologies, provides an integrated information access and management service, and will be marketed to specific industries, as well as consumers and business users. The third generation standards are currently under development under the auspices of the ITU-R (International Telecommunications Union - Radio).

This paper is organized as follows. In Section 2, we discuss some technologies which address the ever increasing need for larger system capacity to serve both high data rates and many simultaneous users, and focus on the concept of transmit diversity. In Section 3, we review some of the service concepts made possible by this capacity, and focus on the need for context-sensitive applications.

2. Technology issues - need for high capacity

2.1. Third generation system technologies

The ITU-R conducts workshops to receive and discuss various proposals for design approaches to IMT-2000. Here we review some proposals presented in Toronto in September 1997 [1][19]. The FRAMES Multiple Access (FMA) concept combines CDMA and TDMA techniques into one harmonized platform also harmonized with GSM. Thus to build multi-mode terminals, GSM and FMA clock rates should be related, and carriers should be placed on a common 200 KHz frequency grid. FMA is proposed in three forms: FMA1 (TDMA without spreading and TMDA/CDMA with spreading) with bandwidth 1.6 MHz and FMA2 (wideband CDMA) with 5 MHz bandwidth, and both Frequency Division Duplex (FDD) and Time Division Duplex (TDD) operation. TDD is included to adapt to asymmetric services (different rates on uplink and downlink) Multiple user data rates are supported in all three forms by multislot, multicode, and variable spreading/multicode respectively. The link can be adapted to changing service requirements, interference, channel state and system loading.

There are 4 different concepts for wideband CDMA, but many of the key parameters have been agreed, e.g. basic chip rate of 4.096 Mchips/s, 10 ms frame length, coherent detection with pilot symbols, and variable spreading for low and medium bit rates, multicode for high bit rates. FMA2 includes packet mode operation where small infrequent packets are attached to a random access request, whereas large and frequent packets are sent using a dedicated channel. FMA2 also proposes support for multi-user detection, interference cancellation and adaptive antenna arrays for both the uplink and downlink receivers.

2.2. Transmit diversity

The first premise of this paper is that one method that can significantly increase system capacity, but which may not have been considered in the working papers for [1] is the idea of using multiple transmit antennas (transmit diversity). For fixed bandwidth and total radiated power, with Rayleigh fading, a wireless system with multiple antennas at both transmitter and receiver has a channel capacity which grows linearly (rather than logarithmically) with the number of antennas [2]. This result, which assumes separate information is sent out of each transmit antenna, holds true even though the transmitter does not know the complex channel transfer characteristic. By explicitly spreading out the antennas well beyond a wavelength, such capacities can be achieved not only on Rayleigh channels with many scatterers, but also on deterministic channels with direct line-of-sight (LOS) paths only and no scatterers [4]. Roughly speaking, the wide spacing replicates the effect of scatterers which create images and thus serves to spread out the apparent source of the signals over a wider angular range. In this way, capacities on the order of $C_{lin} = nlog_2(1+\rho)$ bps/Hz can be obtained for LOS as well as Rayleigh channels when n transmit and n receive antennas are used. In contrast, when the transmit antennas are closely spaced, the number of degrees of freedom on a LOS channel degenerate, resulting in capacities of only $C_{log} = log_2(1+n\rho)$. For other recent work on transmit antenna diversity see [5], [6][9] and [7].

2.3. Capacity expressions

The fundamental result for the capacity in bps/Hz of a wireless system with n_T transmit antennas and n_R receive antennas with an average received SNR ρ (independent of n_T) at each receive antenna was obtained in [2] as

$$C = log_2(det[I_{n_R} + (\rho/n_T)HH^\dagger]) \quad (1)$$

where the normalized channel matrix H contains complex scalars with unity average power loss, and $H^\dagger$ is the complex conjugate transpose of H. The capacity is expressed in bps/Hz in the narrowband limit with no frequency dependence. Normalization is achieved by dividing out the free space power loss and setting the parameter ρ to the desired SNR [1]. This result assumes that H is unknown to the receiver but n_R and ρ are known [2][3]. The signal from each antenna is different. The transmitted data has been split into substreams which are separately independently coded and modulated, and cycled among the antennas to ensure that over time each substream experiences a similar propagation environment.

We consider an environment with only LOS propagation and T and R arrays of $n_T = n_R = n$ antennas. Here we designate the base and subscriber ends of the link as T and R, respectively, but reciprocity applies, and all subsequent results apply to both the downlink (base-to-subscriber) and the uplink (subscriber-to-base).

By replacing $HH^\dagger$ by its singular value decomposition $U\Lambda U^\dagger$ where Λ is a diagonal matrix of eigenvalues and the columns of U are the corresponding unit-magnitude eigenvectors of $HH^\dagger$, we can show

$$C = log_2(\prod_{i=1}^{n}(1 + (\rho/n_T)\Lambda_i) \quad (2)$$

Thus C is maximized when all eigenvalues Λ_i are equal (since H is normalized, the eigenvalues will all be of magnitude 1).

[1]This avoids the need to compute absolute propagation loss and then set the transmitted power to obtain the desired SNR.

If the matrix elements H_{ij} are random variables, then C is also a random variable. In this case, we define an outage threshold x (say 0.01), and define C_x to be that capacity for which $Prob\{C > C_x\} = 1 - x$. For a Rayleigh channel model, we define $H_{Rayleigh}$ as a matrix of normalized (unit magnitude) uncorrelated complex gaussian variates. If $H = H_{Rayleigh}$, then from [3], in the limit of large n and large ρ, the capacity lower bound approaches

$$C = nlog_2(\rho/e) \tag{3}$$

where $e = 2.718...$. This result is close to the maximum capacity C_{lin}, and thus implies that there is a high probability that a particular realization of $H_{Rayleigh}$ has close to n significant (non-negligible) eigenvalues. This result is as expected, since in the Rayleigh model $H_{Rayleigh}$ using *independent* complex gaussian variates, it is implicit that signals arrive from many directions [2] i.e. the apparent transmitter locations are widely spaced.

To find the maximum capacity which may be realized in practice in a LOS environment (no Rayleigh fading), we seek to construct geometric arrangements of the T and R arrays such that all Λ_i in (2) are non-negligible.

To illustrate, we consider a 3-element transmit and receive array, where the R array is linear with interelement spacing $y_r = \lambda/2$, and the T array comprises 3 elements evenly spaced around a circle (e.g. 120 degree sector antennas at 3 separate cell sites in a conventional hexagonal cell layout, all facing inward [10]. For the R array at the center of the cell,

$$H_{LOS} = \begin{pmatrix} e^{-j\theta_1} & 1 & e^{j\theta_1} \\ 1 & 1 & 1 \\ e^{j\theta_1} & 1 & e^{-j\theta_1} \end{pmatrix} \tag{4}$$

can be used to evaluate the capacity at this point. From a histogram of the capacity for all positions in the cell, assuming a path loss exponent of either 2 or 4, we find that for most locations, the capacity is near $C_{lin} = nlog_2(1 + \rho)$ for $n = 3$, so that the benefit of transmit diversity is realized.

2.4. Reduced complexity implementation

To realize the high capacity from transmit diversity requires the processing of signals from a n-element signal vector, as outlined in [3]. To reduce the complexity, the received signal may be premultiplied by a transformation matrix W having fewer rows than columns, thus reducing the length of the signal vector to be processed. After transformation, the capacity

$$C = log_2 \frac{det[WW^\dagger + (\rho/n_T)WHH^\dagger W^\dagger]}{det[WW^\dagger]} \tag{5}$$

If W is square and unitary, then C is unaffected. By choosing $W = \sqrt{n}U_1$ where $U_1 = U$ but with $k < n$ columns, then the capacity is given by (2) with sum to k instead of n. Thus if the eigenvalues beyond k are negligible, then we can use W to reduce the processing complexity, while still retaining most of the capacity.

2.5. Practical implications

Capacities approaching $C_{lin} = nlog_2(1 + \rho)$ can be achieved in a deterministic LOS (non-Rayleigh) environment. This is achieved by spreading out the elements of T either explicitly (by placing one element of T at each of n sites), or implicitly (by adding reflectors which create images of T). This result suggests that in the absence of reflectors, the T array can be explictly spread apart, thus duplicating the effect of images.

The increase in capacity resulting from transmit diversity results in a significant benefit for services requiring high speed data.

3. Service issues - need for sensitivity to context

3.1. Market Demands for services

'Anyplace, anytime communications with anyone' is a common expression of the service concept for future mobile communications systems. However, it may be very difficult to anticipate exactly what kind of services the marketplace will demand and accept [16]. Service concepts proposed for IMT-2000 may be divided into two broad classes: extensions and refinement of existing second generation services, and entirely new services.

The primary existing service of second generation systems is plain old (wireless) telephony, which may be extended with call management features such as call forwarding, voice mail, call waiting, 3-way calling and automatic call handling according to user preferences depending on the caller ID. Another existing service is two-way text messaging (short message service), an extension of one-way paging towards wireless email service.

New services and capabilities for IMT-2000 include [18] provision for new audio, video and data services, including packet data and multimedia services, high service quality and integrity comparable to the fixed (wired) network, bandwidth on demand to support data rates ranging from low speed for paging to high speed for video or file transfer, support for asymmetrical data with high speed in one direction and low speed in the other, security and ease of operation.

[2] For a given interelement spacing at R, see [8] for the angular spread of signals from T required in order to to achieve zero correlation between R elements.

Wireless multimedia services are based on some kind of one or two way video, high speed internet access and web browsing capability. These service concepts arise from the convergence of broadband access and wireless technologies. However, genuinely useful service concepts using these capabilities remains to be specified and tested in the marketplace.

User demands will inevitably drive the evolution of new service concepts. The key demand will be for immediate, *integrated information access and management*, which results from the information explosion, mostly available on corporate intranets as well as the Internet. Since the information exists, users will expect it to be available 'anywhere, anytime', and thus will demand that their portable wireless device has the same power as their office desktop. Users will demand multiple types of wireless access devices, ranging from a small cellphone to a large-screen laptop.

A convenient user interface for all device types is critical to the evolution and acceptance of new service concepts. The challenge is to develop user interfaces which can offer the simplicity of a desk telephone and fax machine and still meet the more complex needs of integrated information access and management. Web browsing using search engines is too complicated for mobile users.

It is useful to be reminded of service concepts that were not successful, such as the videophone [16] or that had only limited success, such as the personal digital assistant (PDA). In the first case, the service did not meet a real need and in fact was considered an invasion of privacy by some. In the second case, the PDA functionality and user interface have not been sufficiently complete and convenient to provide the intended 'assistant' service.

3.2. Context-sensitive service concepts

The second premise of this paper is that a key factor to the success and acceptability of a new wireless service concept is that it be *context- sensitive*, i.e. aware of the circumstances (state) of the user, and sensitive to the anticipated needs when in this state. Perhaps the best context-sensitive service is that provided by a human secretary or butler backed up with instant access to the desktop computer. A practical (non-human) approximation may be provided by autonomous agents [13][12] working with a wireless access device. However, a complete and convenient user interface for the device is essential. An interface which is sufficient for the desktop (e.g. type-point-click), may not be practical for a portable device.

Context-sensitive services may be particularly effective for group projects where there is awareness of the state of all users. Such services may also help to reduce the demand for capacity on the wireless system by 'taking care of things' without requiring too much data transfer between the network and the wireless access device.

A simple example of a context-sensitive service is the automatic configuration, upon entry, of the user's current environment (e.g. temperature, background music) to accommodate user preferences [11], i.e. to create a personal space, in hotel rooms or shared offices.

An instructive example of a context-sensitive service is a tourist (or real estate) application on a PDA, where the tourist is provided with text information about buildings and other landmarks in the city depending on his/her current location and orientation [15]. This example may be implemented using a software form of Post-it notes represented in Standard Generic Markup Language (SGML) [3]. using tags such as <at>(location), <with> (person or object), <during> (time), <facing> (orientation). The software Post-it note consists of the context (tag) plus a body (text), and a collection of such notes comprises a complete document. Notes are triggered dynamically depending on the current context. A related idea is ConTexts (Conversational Texts) [14] in which a complete text is dynamically assembled from a collection of 'chunks' which are triggered depending on the context (parameters) chosen by the reader (e.g. expertise level, amount of detail, language). Each chunk (equivalent to the software Post-it note) is tagged with the relevant parameters. In both cases, information (text chunks or a complete document) is accessed as appropriate depending on the desired context.

Ideally, using a Context-sensitive information access and management service which is aware of the user's current state will be like having a secretary/butler available at all times to respond to the user's current as well as anticipated information needs, as well as to prompt/remind the user of tasks to be done or items requiring a user decision.

4. Summary

In this paper, we have pointed out one technical and one service concept thought to be useful for 3rd generation wireless systems, but which may not be referred to in current IMT-2000 documents. *Transmit diversity* may be used to significantly increase the capacity (bits per second per unit bandwidth) of wireless systems, thus enabling the high data rates needed for integrated information access and management and desktop functionality while mobile. *Context-sensitive applications* may be used to significantly ease the task of information access and management.

[3]HTML, Hypertext Markup Language used to represent web pages is one instantiation of SGML.

Acknowledgements

The author thanks G.J. Foschini for discussions of [2][3], R.L. Kirlin for suggesting the transformation W, S. Catreux for deriving (5) and the figures in Section 2, and m.c. schraefel for the key ideas in Section 3.

References

[1] www.itu.int/imt/

[2] G.J.Foschini, M.J. Gans, 'On limits of wireless communication in a fading environment when using multiple antennas', *Wireless Personal Communications*, accepted for publication.

[3] G.J.Foschini, 'Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas', *Bell Labs Technical Journal*, accepted for publication.

[4] P.F. Driessen, G.J. Foschini, 'Deploying multiple transmit antennas to greatly enhance capacity in wireless systems', *IEEE Trans. Comm.*, under review.

[5] V. Weerackody, 'Characteristics of a simulated fast fading indoor radio Channel," *Proceedings 43rd IEEE Vehicular Technology Conference*, pp. 231-235, Secaucus, NJ, May 1993.

[6] V. Weerackody, 'Diversity for the Direct-Sequence Spread Spectrum System Using Multiple Transmit Antennas', *Proceedings IEEE International Conference on Communications*, pp. 1775-1779. Geneva, Switzerland, May 1993

[7] N. Seshadri and J. H. Winters, Two Signalling Schemes for Improving the Error Performance of Frequency-Division-Duplex (FDD) Transmission Systems Using transmitter Antenna Diversity, *International Journal of Wireless Information Networks*, Vol. 1, No. 1, 1994, pp 49-60.

[8] J.Salz, J.H.Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio", *IEEE Trans. Vehic. Tech.*, vol. 43, no. 4, pp. 1049-1057, November 1994.

[9] S. Ariyavisitakul, T. E. Darcie, L. J. Greenstein, M. R. Phillips and N. K. Shankaranarayanan ' Performance of simulcast wireless techniques for Personal Communication Systems', *IEEE Sel. Areas Commun.*, Vol. 14, No. 4, May 1996, pp. 632-643.

[10] F. Borgonovo, A. Acampora, 'Capture division multiple access: a new cellular access architecture for future PCN's', *IEEE Comm. Mag.*, Vol.34, No.9, September 1996, p. 154-162.

[11] Bill Gates, *The Road Ahead*, Microsoft Press 1996.

[12] K. Decker, A. Pannu, K. Sycara, M. Williamson, "Designing behaviours for information agents", *Proceedings of the First International Conference on Autonomous Agents*, Marina del Ray, CA, February 1997, ACM Press, 197, pp. 404-412.

[13] M.N. Huhns, M.P Singh, "Conversational Agents", *IEEE Internet Computing Magazine*, March-April 1997, pp. 73-75.

[14] m.c. schraefel, "ConTexts, Intensional Document Creation, Delivery and Retrieval", *Proc. IEEE Pacific Rim Conference on Cmmunications, Computers and Signal Processing*, Victoria, BC August 1997, pp. 417-419.

[15] P.J. Brown, J.D. Bovey, X. Chen, "Context-Aware Applications: From the Laboratory to the Marketplace", *IEEE Personal Communications Magazine*, Vol. 4, No. 5, October 1997, pp. 58-64.

[16] R.W. Lucky, "New Communications Services - What Does Society Want?", *Proc. IEEE*, Vol. 85, No. 10, October 1997, pp. 1536- 1543.

[17] M. Shafi, et al, "Wireless Communications in the Twenty-First Century: A Perspective, *Proc. IEEE*, Vol. 85, No. 10, October 1997, pp. 1622-1638.

[18] K. Buchanan et al, "IMT-2000: Service Provider's Perspective", *IEEE Personal Communications Magazine*, Vol. 4, No. 4, August 1997, pp. 8-13.

[19] R.D. Carsello et al, "IMT-2000 Standards: Radio Aspects" *IEEE Personal Communications Magazine*, Vol. 4, No. 4, August 1997, pp. 30-40.