1
How Can Anyone Afford Mobile Wireless Mass Media Content?

Timothy X. Brown

1 Overview

This chapter examines the cost of delivering multi-media content-on-demand to end users via a wireless media. The analysis lays out the technological and spectrum constraints that will shape any widespread deployment of wireless multi-media applications. It considers only existing and near-term delivery technologies. It does not address customer demand, the end-user terminal, nor content specific issues such as content generation, cost of content, billing models, and copyright. It takes as a baseline the current second generation wireless cellular architecture in the U.S. Three representative contents are analyzed: real-time video, audio, and text news. The analysis suggests that even low-fidelity high-compression video and audio will remain expensive at \$0.12–\$0.28 per minute. Higher fidelity would require a significant increase in bandwidth and would be five to 40 times more expensive. Mainly text news, on the other hand, requires little bandwidth and would be inexpensive.

Various technological advances, such as one-way only transmission, third generation wireless, and non-real-time delivery, could collectively reduce the cost by a factor of approximately 8. Significantly more savings are possible in a broadcast-like model that would share the delivery cost among many users. Alternatively, WLAN-based architectures that restrict access to Internet-like content downloads in localized hot-spots are an affordable delivery model. Therefore, while a brute force on-demand, anywhere, high-quality real-time content model may not be viable, alternative delivery models may prove quite economical.

The chapter begins by developing a simple technology-based economic model for delivering content via current cellular technology. It then examines alternatives for reducing the delivery cost.

2 The Economic Challenge

In this section we will establish that under current wireless deployments, individual on-demand high-quality audio or video delivered to wireless mobile users would be prohibitively expensive. Further we will establish baseline technology parameters which will clarify the driving factors behind the cost and be the starting point for our discussion of potential solutions to this problem in the next section.

Our strategy is to estimate the price charged by current wireless cellular providers to deliver a given bit rate, the so called *bit rate price.* We make a number of rather gross assumptions in this analysis, but, as we will see, even, if the price estimates are off by a factor of two or three, the conclusions would remain unchanged.

Wireless cellular in the U.S. is quite competitive with up to nine different service providers in every major market.[1] With this level of competition, we assume that profits are limited and the price approximates the cost of providing the service. The cost of providing the mobile telephone service, as we will establish in estimating the bit rate price, is, to a first order, proportional to the bit rate of the service. Therefore, by comparing the bit rates of different services to the bit rate price, we can estimate the price at which these services would be offered.

2.1 Estimating the Bit Rate Price

The price per minute of a mobile telephone call divided by the data bit rate of the telephone call yields a price per bit rate for the phone call. In this section we will estimate the price per minute and bit rate for voice, and then show that the resulting price per bit rate applies to other mobile communication services at other bit rates.

The price per minute of use (MOU) for a mobile telephone call could be estimated through the many different service pricing plans. Unfortunately, these plans do not correctly state the price per MOU. For instance, if someone pays $0.10 per MOU for 300 MOU and only uses 150 minutes, then their effective price per MOU is $0.20. Further, prices change between peak and off peak times. We do not have the data to find true prices and usage times for individuals. In particular, we do not have data limited to peak demand periods that would best reflect the marginal cost of providing a minute of service. However, we can still estimate a price per MOU through aggregate statistics. A survey of consumer data results in an FCC study (FCC, 2003) yields an average monthly MOU that has risen from 380 minutes in 2001 to 427 minutes in 2003. The average subscriber monthly bill has risen from \$47.37 in 2001 to \$48.40 in 2002. The total bill divided by the monthly MOU is an upper bound on the marginal price of one MOU since the monthly bill includes fixed cost such as billing. If instead we look at the yearly increase in MOU and the yearly increase in the monthly bill we can derive a lower bound estimate on the marginal price of an MOU. This is a lower bound since price per MOU is falling, and tends to depress the overall monthly bill. From the above data, we derive the following estimate on the price per MOU:

$$\$0.11 = \frac{\$48.40}{427\text{MOU}} > \text{PriceperMOU} > \frac{\$48.40 - \$47.37}{(427-380)\text{MOU}} = \$0.022 \qquad (1)$$

Next we consider the bit rate of a mobile telephone call. In the U.S., most mobile phones are digital and based on one of three air interface technologies, the so-called USDC (IS-54), CDMA (IS-95), or GSM. Other digital technologies, such as PACS and iDen represent less than 10% of the total market and so are not considered. Next generation standards will be discussed later.

Table 1 shows the raw channel bit rate, the raw per user bit rate after removing control, synchronization, and signaling overhead, and the final usable data rate after removing error correcting coding overhead.[2] We don't consider any additional overhead specific to multimedia delivery (Radha et al., 2000).

Table 1: Various bit rates of 2G standards (Rappaport, 2002)

standard	raw channel bit rate (kbps)	per user bit rate (kbps)	usable user bit rate (kbps)
GSM	270.8	22.8	13.0
USDC	48.6	13.0	7.95
CDMA	1228.8	14.4	6.65

The only relevant rate here is the usable user bit rate (the others are included to suggest the overhead that is necessary to enable a high-quality wide-area high-mobility radio connection). This shows a range of 6.65kbps to 13kbps for these different standards. Combining these values with (1) we come up with the following bound.

$$\$0.017/\text{kbps} = \frac{\$0.11}{6.65\text{kpbs}} > \text{price per bit rate} > \frac{\$0.022}{13.0\text{kbps}} = \$0.0017/\text{kbps}$$

We are only interested in a gross estimate so we simplify these bounds to:

$$\begin{aligned}\text{Price per minute per bit rate} = P_b &= \$0.01/(\text{min kbps}) \\ &= \$10.00/(\text{min Mbps})\end{aligned} \quad (2)$$

The result in (2) can be interpreted to say that for typical cellular and PCS deployments in the U.S., the cost to deliver a reliable 10kbps data rate to a user is $0.10 per minute. It should be realized that this cost is a cost per volume of data, that is:

$$\frac{\$0.10}{\text{min}\,10\text{kbps}} \times \frac{1\,\text{min}}{60\,\text{sec}} \times \frac{8\text{bit}}{\text{byte}} \times \frac{100\text{kB}}{\text{MB}} = \$1.33/\text{MB} \quad (3)$$

In some cases, such as pricing news delivery (i.e., downloading a known size file), we will work directly with this volume price, and not concern ourselves with the delivery speed.

2.2 Estimating the Service Price

The result in (2) can be further extended to show that it scales over a range of bit rates, that is, $P(B)$, the price to offer a service at a given bit rate, B, is simply:

$$P(B) = P_b\,B. \quad (4)$$

By estimating the bit rate of different services, we can estimate the price of offering these services.

In so-called 2.5G standards (such as the IS-54 variant, IS-136, or the GSM variant GPRS) users can receive different data rates. IS-136 is a time division multiple access (TDMA) scheme with 6 time slots per frame. A normal user receives two time slots per frame to yield a 7.95kbps usable data rate. A user who wanted half the rate would use only one time slot, and a user who wanted 3 times the data rate would use all 6 slots. GPRS which is also TDMA behaves similarly, i.e. a user who doubles their data rate doubles the resources taken from the network (Rappaport, 2002). In IS-95, users can use greater or lower bit rates, and as in the TDMA system, a user consumes communication resources in proportion to the bit rate (Gilhousen et al., 1991).

If users double their bit rate they would require twice the capacity resources. As argued earlier, mobile telephone service has a high degree of competition in the U.S.. Further, capacity in terms of base stations is increasing at 20% per year (Roche et al., 2004) in order to keep up with demand. There is no reserve of excess capacity to accommodate higher data rate customers. Therefore, in order to cover their costs, service providers would need to double their price to users with twice the bit rate as indicated in (4).

This argument does not extend indefinitely. For instance, very low data rate services like short messaging services use a different packet switched-like model for delivery. At very high data rates we exceed the maximum bit rate per channel of the standard. For our purposes though, we assume that the price scales linearly with required bit rate as needed. For instance, very high data rates could be accommodated by using multiple channels.[3]

Based on this notion, what are the costs of providing our baseline real-time services of high-quality stereo audio and high-quality full-frame video? The bit rate of near CD-quality MPEG layer 3 (MP3) audio is 128kbps (Puri et al., 2000a). Similarly for television quality MPEG 1 video with stereo sound, the data rate is 1.2Mbps. For news, the New York Times home page is approximately 150kB.[4] Applying these to (3) and (4), yields $1.28/min for audio, $12.00/min for video, and $0.20/download for news.

What do these prices mean? Watching a 100 minute movie would cost $120. Listening to a single three-minute song would cost $3.84. Downloading 10 news stories in the process of reading the morning news would cost $2.00. This is above and beyond the price of the content. It is unlikely that content providers would agree to a revenue split that is so heavily weighted toward content delivery. So, it is unlikely that content would be available even if some segment of consumers were willing to pay these prices.

Clearly these prices must be significantly reduced to achieve mass market acceptance. So, what can be done? The basic problem is that the bit rate demand is high compared to current providers capacity. So broadly, we can consider two approaches: lowering the user bit rate demand or lowering the price by increasing capacity resources. In other words, in (4) we can reduce B or we can reduce P_b. The next two sections discuss each of these in turn.

3 Reducing the Bit Rate

This section describes several methods of reducing the effective bit rate from which the reduction in the price of the service can be calculated directly . The methods considered include lower bit rate encoding, one-way vs. two-way communication, and better channel spectrum efficiency.

Lower Bit Rate Encoding: It is difficult to reduce the bit rate through encoding while preserving the content quality. Current encoding of audio and video content as well as the images that make up the bulk of web page downloads is relatively compact.

Better high-quality coders are an active area of current research. For instance, the H.26L standard is being developed to work at 800kbps and research shows that rates including high quality stereo audio could be pushed to 400kbps (Dumitras & Haskill, 2002). The lowest of such rates are computationally intensive and can not be encoded in real time. This suggests that we might expect a factor of two reduction in the bit rate in the near term.

One-way vs. Two-way: Another way to reduce the effective bit rate is to use the asymmetry in the communication. The price per bit rate has assumed a symmetric two-way up and down link connection. But audio and video content is typically unidirectional and therefore uses one half of the spectrum. This can reduce the service price by half.

One difficulty in implementing such a strategy is that the current mobile cellular requires a two-way connection for connection maintenance. Mobiles assist in handoffs, acknowledge system commands, and so on. This overhead is necessary for connection quality and could not be completely removed; reducing the possible savings.

Better Spectrum Efficiency: Another approach is to use different radio techniques which in effect increase the number of mobile telephone channels per base station. For instance, if a new radio technology were deployed that could double the number of channels per base station at no extra cost, then we could halve the cost per channel. Since much of the cost is bound up in the base station, spectrum, and backhaul, the cost of a new radio technology would be a minority factor in the overall cost.

Unfortunately, we do not foresee any radio technology with significant spectrum efficiency improvements in the near term. The most prominent technology development is so-called third generation wireless (3G). CDMA2000 is the next generation of IS-95. Currently CSDMA2000 has two versions being deployed, CDMA2000 1X which is similar to IS-95, but allows shared user rates up to 144kbps; and CDMA2000 EV-DO which has shared user rates up to 2,400kbps. Early tests suggest effective rates for the EV-DO version of 400–800kbps (Seybold, 2002). Another 3G technology, WCDMA, is promoted by GSM carriers. It has rates similar to CDMA 2000 EV-DO (although using 3 times the spectrum). WCDMA is not expected to have widespread deployment for five years (Seybold, 2002). Third generation wireless technologies are primarily focused on providing a wide variety of data rates and services and are projected to have only a factor of two improvement in spectrum efficiency (Nettleton, 2002; Seybold, 2002). Recent pricing announcements suggest that prices for newer technology will be close to (3). For instance, AT&T Wireless prices 1MB of data at $1 for high volume users (ATT, 2004).

Other technologies on the horizon such as so-called 4G and multibeam antennas have promise to provide large efficiency gains. But they are considered to be still 5 to 10 years out and so outside the scope of this chapter (Gitlin, 2002). Satellites provide mobile service. Generally, service prices are higher than cellular. Satellites are not a solution to the economic challenge.

Combining all the near-term advances suggested in this section, we might expect a factor of 8 reduction in the cost of delivering the multimedia content. These advances suggest a cost of $0.16/min for audio, $1.50/min for video, and $0.025 per page of news.

4 Increasing Capacity

The bit rate price is affected by limits in spectrum and capacity for providing service. If the capacity could be increased cost-effectively, the bit rate price may be able to be reduced. This section looks at ways to increase the effective capacity. It analyzes the availability of more licensed spectrum, the potential for using unlicensed spectrum, and the cost of more base stations.

4.1 More Licensed Spectrum

Currently the U.S. has 180MHz of spectrum for mobile telephony; 50MHz in the 800MHz cellular band, 120MHz in the 1900MHz PCS band, and at least 10MHz in the 800MHz ESMR band. This section explores other bands that have potential for offering mobile communication services both now and in the future.

The first constraint is allowed frequencies. Frequencies above approximately 3000MHz are unsuitable for a mobile wireless application because they require line-of-site or near line-of-site to maintain a connection. Low frequencies below approximately 100MHz are not suitable since efficient antenna sizes become unwieldy for a low power mobile device.

Between 100 and 3000MHz other licensed bands include the MDS bands at various frequencies between 2.150 and 2.680 GHz (FCC, 2001a), and the WCS bands at 2.3GHz (FCC, 2001b). While these bands constitute a large amount of spectrum (108MHz), they are limited to fixed and not mobile applications. The significance of these limitations is reflected in the bidding which resulted in only $200M total from operators.

The FCC is licensing more spectrum for mobile applications via spectrum auctions. Potential spectrum here includes the so-called lower and upper 700MHz Bands. These have 78MHz of bandwidth between them and are located in an ideal band with respect to propagation and radio equipment. They are currently encumbered with existing TV stations especially in urban markets, but these stations are expected to migrate to lower frequency channels by 2006 (FCC, 1997).

The FCC and NTIA have proposed sharing plans for both the MDS bands and other bands currently controlled by the federal government. Under the rules, 3G-like services could be offered which meet certain interference limitations with the current incumbents. These limitations are so onerous that it is not believed that any major metropolitan market would be able to offer 3G service with this band (Weingardt & Murphy, 2001).

The recent C and F block broadband PCS auctions generated $16.9B for up to 30MHz of spectrum in each market[5] (FCC, 2001d). The bulk of the revenue was generated in the densest urban markets. For instance, the 30MHz of licenses for New York City generated $5.6B ($300/pop)[6] and similarly Los Angeles generated $1.5B ($200/pop). The 30MHz in these two auctions represents a 20% increase (150MHz to 180MHz) in available spectrum. These bids suggest significant increases in mobile wireless spectrum must be made before spectrum demand becomes saturated. This suggests that new licensed spectrum is an expensive way to add spectrum, on the order of building more base station infrastructure as described later.

4.2 Unlicensed Spectrum

Unlicensed spectrum provides an opportunity for expanding the available spectrum without the costs associated with the spectrum auctions. In the U.S., there are three prime unlicensed bands in the 100 to 3000MHz range with significant bandwidth: The 902 to 928MHz ISM band, 1.91 to 1.93GHz PCS bands, and the 2.4 to 2.4835GHz ISM band (FCC, 2001d).

The 900MHz band is used by cordless phones and was used by Metricom to offer their Ricochet service. Power limits for the band limit the effective range. The PCS unlicensed band is the smallest band and has used restrictions that would limit widespread use of the band.

The 2.4GHz is a large band which has proved quite successful for wireless local area networks (WLAN). Some operators are extending this model to provide fixed wireless Internet access. Wide-area Internet access is threatened in the long term as increasing interference from more users reduces the ability to make long connections. The more promising application is for providing hot spot access at hotels, airports, restaurants, etc. This relies on short range connectivity using a plethora of low-cost 802.11 access points. We will expand on this idea later.

4.3 More Base Stations

The wireless cellular concept allows operators to increase capacity within a fixed amount of spectrum by building more and more base stations (Rappaport, 2002). For a given amount of spectrum, each base station has a fixed quantity of capacity. The capacity of an operator's system is proportional to the number of base stations.

Table 2 shows the U.S. mobile telephone industry's annual capital investment, increase in number of base stations, increase in number of subscribers, investment per added base station, and investment per net added subscriber for the past 10 years. The total investment per base station has generally held close to $1M per base station for a decade. A look at investment per subscriber tells a similar story where investment per added subscriber has fluctuated around $900. Over this period, the cost of adding new capacity has been constant (to within a factor of 2).

Table 2: Mobile Telephone Industry Annual Added Cell Site and Subscriber Costs for the U.S. 1991–2001 (Roche et al., 2004)

year	investment (millions)	increase in cell sites	increase in subscribers ('000)	investment per cell site	investment per added subscriber
1994	$4,983	5,096	8,125	$978,000	$613
1995	$5,141	4,743	9,652	$1,084,000	$532
1996	$8,494	7,382	10,260	$1,151,000	$828
1997	$13,480	21,560	11,270	$625,000	$1,196
1998	$14,480	14,290	13,900	$1,013,000	$1,042
1999	$10,720	15,810	16,840	$678,000	$637
2000	$18,360	22,590	23,430	$813,000	$784
2001	$15,410	23,250	18,896	$663,000	$815
2002	$21,900	11,800	12,392	$1,856,000	$1,767
2003	$18,940	23,650	17,955	$801,000	$1,055

The $37B cost of spectrum licenses adds a fixed cost shared by every base station and every customer. June 2004 had an estimated 148M subscribers and 150,000 cell sites. These numbers yield an added $250/subscriber and $250,000 per cell site. Both of these numbers are smaller than the capital outlay number and will decrease with continued growth. While significant, the added spectrum cost does not change the conclusion that the cost per cell site or subscriber will not decrease significantly in the near future.

In this section, we have shown that adding spectrum or base stations are expensive propositions that will not likely reduce the bit-rate price.

5 Quality

So far we have considered the cost of delivering on-demand, anywhere, high-quality real-time content. This section and the next three sections consider the gains from relaxing these requirements along each of the three dimensions: quality, choice, location. First we consider the role of high-quality content.

Multimedia content can trade off quality for lower bit rates in a very direct way. Audio can be reduced to very low rates. Mobile telephony encoding speech (but not general audio such as music) at approximately 10kbps has been the starting point of this analysis. General audio rates between 12 and 256kbps can be achieved where 12kbps is for near telephone quality mono and 265kbps is for high fidelity stereo (Audioactive, 2002; Bouvigne, 1998).

Lower quality video encoders reduce the bandwidth by reducing the screen resolution, reducing the number of frames per second, or allowing more encoding artifacts to be introduced into the screen. For instance, video conference quality is 384kbps and a graing thumbnail is 28kbps (Bhaskaran & Konstantinides, 1997; Puri et al., 2000b).

For news, the main content is text. But much of news web pages are made up of ads and figures. The text content of the New York Times home page and top news stories consists of approximately 6kB of text per download.

The lowest bit rates are 28kbps for video, 12kbps for audio, and 6kB/download for news. Combining these rates with the near-term advances in reducing the bit rate price (the factor of two for one-way transmission, and the factor of two for better spectrum efficiency), these translate into $0.07/min, $0.03/min, and $0.002/download.

The results in this section show text news can be quite cost effective at a fraction of a cent per download. Audio and video can be brought to below current prices per minute for cellular voice. But, the quality of the telephonegrade audio and thumbnail video is so low that it is not likely consumers would be willing to pay anything. The higher quality content has 4 to 40 times higher bit rates, and would have as a result prices which are again too expensive.

6 Location

The delivery model has so far assumed that consumers could download content anywhere they can make a mobile phone call; inside or outside, urban or rural areas, standing still or moving at high-speeds. In this section we consider a model suitable only for stationary users at localized coverage locations.

WLAN's are being deployed at many levels by end users, private service providers, free service providers, and paid service providers. We will focus on the WLAN standard which is by far the most popular, 802.11b also known as WiFi.

WiFi networks are spreading rapidly. Over 40% of large corporations were using wireless networks at year end 2002. End users are buying WiFi interface cards in great numbers. Laptop computers come with integrated WiFi interfaces. Further, many are installing WiFi networks within their homes (Werbach, 2001).

Private service providers include campuses and offices which offer the service to a limited proscribed community such as students or employees. Many organizations, especially in the high-tech sector, are deploying such networks in this manner.

Free service providers include restaurants, individuals, and some organizations that choose not to restrict access or charge for providing connectivity. Many public spaces have service provided in this way. Perhaps the most interesting phenomenon here is the emergence of "community networks" consisting of individuals who make their wireline access available to all.

Finally, paid service providers sell WiFi-based access to subscribers. Often these services are deployed in airports, business hotels, and the like, catering to business travelers.

The main point here to note is that operators do not pay for their spectrum, equipment is cheap, and the marginal impact of additional users on low cost DSL and cable modems is negligible. As a comparison Verizon Wireless offers broadband access using 3G technology for $80 per month. T-Mobile offers a similar service at WiFi hotspots for $30 per month.

The unlicensed spectrum is subject to unregulated interference but, if intended coverage is kept localized around the access points, radio link budgets can retain a sufficient interference margin. In other words, the opportunity cost to existing WLAN operators for a new WLAN access point is negligible if the coverage of the new access point is localized. The net effect is that each additional access point adds more capacity similar to adding base stations in cellular.

WLAN cards are under $100 and access points are less than $300.[7] Access point installation often consists of attachment to a wall or ceiling and running a cable to the network interface. Some vendors sell integrated wireless access points and DSL/cable modems. The cost of high-speed access for heavy business users is $300/month.

To put these costs in perspective, suppose that we want nationwide coverage which we define as one access point for every 20 persons in the U.S. This yields approximately 13 million access points. Conservatively, we estimate a $1000 cost for each access point and installation and each access point requires a single high-speed access connection. This yields a $13B installation cost and an operating cost of $47B per year. The installation costs are comparable to the annual wireless capital investment listed in Table 2. The $47B operating cost is comparable to the $88B in wireless cellular revenue from in 2003 (Roche et al., 2004). Unlike the concentrated investment in Table 2, these costs are distributed across many individuals and organizations as suggested above.

To look at it another way, suppose each access point is used during one busy hour each day. Under the above assumptions, the cost of the access point and one year high-speed data service is $4600. If each access point is used for one busy hour per day, this is 22,000 minutes of use per year. Assume that the throughput of the access point and high-speed modem is 1Mbps (conservative given the 11Mbps peak rate). If the $4600 is to be recovered in one year, then the cost per minute of use per bit rate is $0.21/(min Mbps). Comparing this with the result in (2), this suggests a bit rate price that is 50 times lower.

Though significantly cheaper, the service provided is not the same as in the cellular model. The coverage is localized to the hot spots and not wide area. The users must be stationary or have limited mobility during the content delivery. The connection service is more similar to today's internet. In particular, no quality of service guarantees are provided like in the circuit switched (i.e., reserved bandwidth) model of traditional cellular. As in streaming media over the Internet, significant buffering is necessary to account for the more erratic internet performance. Though not a direct substitute for cellular delivery, it may have an important role.

7 Choice

There is in fact a very efficient mechanism for delivering wireless multi-media content: broadcast radio and television. Here the marginal delivery cost per user is zero. In the broadcast extreme, the delivery is efficient, but the content choice is limited in both what can be viewed and when it can be viewed.

7.1 Content Sharing Models

Broadcasting is efficient since every user shares a single transmission. Content on demand on the other hand requires one transmission per user. We will discuss several models intermediate to these alternatives.

One method to increase time flexibility is to broadcast time shifted versions of the same content. For instance, a 90 minute movie could be broadcast on 9 channels each shifted by 10 minutes. A user would never have to wait more than 10 minutes to watch the movie. Further, they could fast forward, rewind, and pause in 10 minute increments. As long as at least 9 users shared the content, it would be a net win in spectrum efficiency.

Another approach is to aggregate content on demand requests. For instance, if two users ask to watch the same movie, then a single stream would serve both. This model could be facilitated by synchronizing starting points. For instance, all movies would start on the hour. As the hour approached, users would submit their requests, and then only one transmission would be required for each distinct request.

A third approach would be to offer many different content alternatives but only broadcast the content actually being viewed. For instance, users would be offered a choice of 1000 different channels, but, in a single cell, if only 10 different channels are actually being viewed, only those 10 are broadcast. More detailed user preference and traffic data would be necessary to show the viability of such a model.

The main principle here is that the cost per user decreases as more users share the content. In particular, the marginal delivery cost is zero after the first user. This suggests that pricing should be designed to encourage content sharing. Potentially popular content should be given a lower price in order to steer users to common content. Less popular content that is less likely to be viewed by multiple users should be more expensive to reflect its true marginal cost.

7.2 Real Time vs. Non-Real Time

In an immediate content-on-demand model subscribers can choose any content and receive it immediately. In the content sharing models, user choice and flexibility are reduced to encourage coincident requests. Alternatively, subscribers could choose a future playback time and the content could download over time as bandwidth is available. This approach would require significant storage and would not be appropriate for all user terminals. Non-real-time delivery can stop and resume during congestion times while real-time delivery requires a continuous commitment from the network once the streaming begins. Real-time content requests must be rejected if bandwidth is not immediately available.

The advantage of non-real-time delivery can be demonstrated with the following model. Requests for downloads arrive over time. The service provider has *C* channels set aside for the multimedia downloads. If the service provider has *C* or fewer requests, they are all streamed out. If more than *C* are active the excess are queued and served in order as active streams finish. Under standard

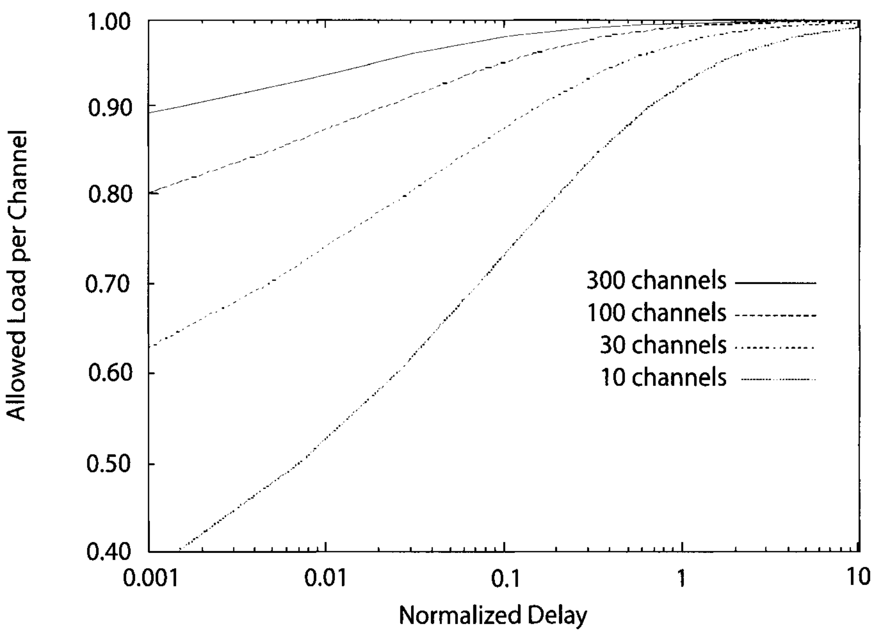

Figure 1: The normalized channel utilisation as a function of the wall time to service (normalized by the average time) for a non-real-time delivery model

traffic assumptions we can compute how much load the system can carry as a function of the average wait time before a request starts service.[8]

Figure 1 plots the channel utilization as a function of delay. For example, let resources for $C = 10$ video sessions be set aside and video sessions consist of 10 minute clips. If clips must be served within 6 sec, normalized average delay of 0.01, then the load can only be about one half. On the other hand, if users can wait 10 minutes (normalized delay of 1) then the load exceeds 90%.

This model shows that with a small number of available channels, utilization approximately doubles. In other words, non-real-time delivery is twice as efficient as real-time delivery. The efficiency increase can be even higher if combined with the content sharing models where common requests are aggregated and sent out as a single stream.

Perhaps more important than the increased efficiency, non-real-time delivery is more suitable for the WLAN model described previously which on average has large low-cost bandwidth, but over time can have significant performance variations which make it unsuitable for real-time delivery. The main drawback to non-real-time delivery is the lack of on-demand downloads and that the mobile device needs sufficient storage for the content.

8 So What Can Be Done?

Unmodified video and audio content on demand will never be economically viable over traditional cellular and PCS networks. Adding more costly spectrum or base stations will not change these economics. Their viability increases dramatically with compression and trading off lower fidelity for lower bandwidth. However, users will be less willing to pay for lower fidelity and it is not clear that a viable price/fidelity combination exists. Greater efficiency, such as with newer cellular protocols, improves this situation, but is not in itself enough. Many current wireless providers have taken this approach; offering low resolution images and video delivery.

For high-fidelity content what are needed are newer models for downloading. Broadly, two approaches appear viable. The first is to move from high-cost cellular to a low-cost WLAN based hot-spot approach to wireless access. The WLAN approach is better suited to a download now and view later model and can be very cost efficient. The other approach is closer to a broadcast model where the content that can be delivered is constrained to encourage content sharing. Both of these approaches remove much of the spontaneity of a content on demand model but appear to be the most promising approach to cost effective multimedia content delivery.

Endnotes

[1] Between the Cellular (2), PCS (6), and ESMR (1) bands any given area in the U.S. has as many as 9 different mobile telephone licenses Year end of 2000 data indicates that 83% of the U.S. population had a choice of 5 or more providers (FCC, 2003).

[2] IS-95 first adds signaling overhead and then adds coding overhead. Further, IS-95 uses a variable rate vocoder. We assume the vocoder works at 50% load on average.

[3] One of the 3G standards, CDMA-2000, uses a multi carrier technology that combines multiple channels for higher-rate users.

[4] Estimated from the size of the www.nytimes.com web-page downloaded on 03/13/02 and 03/14/02.

[5] These auctions offered licenses that were unsold or unpaid for in an earlier auction. So coverage was sporadic.

[6] The license price divided by the population in the licensed area.

[7] The prices in this section are based on a casual survey of various vendors and service providers at the time of writing. They are intended to be suggestive of the costs of WiFi service relative to conventional wide area cellular.

[8] In particular, requests arrive as a Poisson process, the average time to serve a request is exponentially distributed, and as a result the Erlang C distribution applies.

References

AT&T. (2004). Mmode plans. Retrieved September 6, 2004 from the World Wide Web: http://www.attwireless.com/personal/features/mmode/plans.jhtml.

Audioactive. (2002). *Audioactive internet bit rate versus quality (technote).* Retrieved October 3, 2002 from the World Wide Web: http://www.audioactive.com/intro/papers/bitrate.html.

Bassuener, K. (2002). Verizon wireless rolls out new data prices. *Wireless Week.*

Bhaskaran, V. & Konstantinides, K. (1997). *Image and Video Compression Standards: Algorithms and Architectures.* Kluwer, second edition.

Bouvigne, G. (1998). *Listening tests of mp3 using different bitrates.* Retrieved October 3, 2002 from the World Wide Web: http://www.mp3-tech.org/tests/gb/.

Dumitras, A. & Haskill, B. (2002). An encoder-only texture replacement method for effective compression of entertainment movie sequences. In *IEEE Internationsal Conference on Acoustics, Speech, and Signal Processing.* Orlando, Florida.

FCC. (1997). *Fifth report and order: Advanced television systems and their impact upon the existing television broadcast service.* Report and Order 97–116, Federal Communications Commission released April 27.

FCC. (2001a). *Auction 06: Multipoint/multichannel distributions services fact sheet.* Retrieved from the World Wide Web: http://wireless.fcc.gov/auctions/06/, updated August 2.

FCC. (2001b). *Auction 14: Wireless communication service (WCS) fact sheet.* Retrieved from the World Wide Web: http://wireless.fcc.gov/auctions/14/, updated August 2.

FCC. (2001c). *Auction 35: C and F block broadband PCS fact sheet.* Retrieved from the World Wide Web: http://wireless.fcc.gov/auctions/35/, updated September 12.

FCC. (2001d). *Unlicensed radio frequency devices.* Code of Federal Regulations Title 47, Chapter 1, Part 15, Federal Communications Commission. Revised October 1.

FCC. (2003). *Eighth annual report and analysis of competitive market conditions with respect to commercial mobile services.* Report 03-150, Federal Communications Commission. released July 14.

Gilhousen, K., Jacobs, I., Padovani, R., Viterbi, A., Weaver Jr., L., & Wheatley III, C. (1991). On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology,* 40(2):202–211.

Gitlin, R.D. (2002). Challenges and issues for 4G/5G wireless networks. In *IEEE Wireless Communications and Networking Conference.* Orlando FL, March 17–21.

Nettleton, R. (2002). Third-generation cellular: Technical and economic challenges. In *2002 International Symposium on Advanced Radio Technologies.* Boulder, CO, 4–6 March.

Puri, A., Schmidt, R.L., & Haskell, B.G. (2000a). Overview of the MPEG standards. In Puri, A. & Chen, T. (Eds.), *Multimedia Systems, Standards, and Networks,* Chapter 4, (pp. 87–129). Marcel Dekker Pub.

Puri, A., Schmidt, R.L., Luthra, A., Chen, X., & Talluri, R. (2000b). MPEG-4 natural video coding—part I. In Puri, A. and Chen, T. (Eds.), *Multimedia Systems, Standards, and Networks,* Chapter 8 (pp. 205–244). Marcel Dekker Pub.

Radha, H., Ngo, C.Y., Sato, T., & Balakrishnan, M. (2000). Multimedia over wireless. In Puri, A. and Chen, T. (Eds.), *Multimedia Systems, Standards, and Networks,* Chapter 19 (pp. 525–557). Marcel Dekker Pub.

Rappaport, T. (2002). *Wireless Communications Principles and Practice.* Prentice-Hall, second edition.

Roche, R., Jobanputra, P., & Rodriguez, L. (2004). *CTIA's wireless industry indices, semi-annual data survey results, January 1985–December 2003.* Survey summary, Cellular Telephone Industry Association.

Seybold, A. (2002). Wireless Data University Training Seminar, Andrew Seybold Group, LLC. Las Vegas, October 15.

Weingardt, B.H. & Murphy, R. (2001). The search for 3G spectrum: Part II. *Wireless Broadband,* 2(6). June/July.

Werbach, K. (2001). The paradise of the commons. *Release 1.0,* 19(10).