IEEE P 802.20™/PD<insert PD Number>/V<insert version number>

Date: <July 8, 2003>

Draft 802.20 Permanent Document

Traffic Models for IEEE 802.20 MBWA System Simulations

This document is a Draft Permanent Document of IEEE Working Group 802.20. Permanent Documents (PD) are used in facilitating the work of the WG and contain information that provides guidance for the development of 802.20 standards. This document is work in progress and is subject to change.

Contents

Traffic Models for IEEE 802.20 MBWA System Simulations

[Editor's Note: Following are the relevant contributions on this topic so far. Please let me know of any that I have missed. This document is very much a work in progress. So let's have some discussion. While everything in this document is a subject for discussion, some parts are highlighted in particular as a discussion point]

C802.20-03/43 (& 03/57) has a detailed proposal for traffic models for Web-browsing, FTP, WAP, and near real time video.
C802.20-03-13r1 details a user modeling approach including a Web/interactive user/capacity model.
C802.20-03/35 gives a list of MBWA traffic types.
C802.20-03/53 shows a measurement of the mix of traffic types.
C802.20-03/46r1 states that a mix of narrowband of broadband traffic types should be used]

1 Overview

1.1 Purpose

This document specifies a set of mobile broadband wireless traffic models in order to facilitate the MBWA system simulations.

1.2 Scope

The scope of this document is to define the specifications of mobile broadband wireless traffic models.

1.3 Abbreviations and Definitions

FTP = File Transfer Protocol
HTTP = Hypertext Transfer Protocol
MBWA = Mobile Broadband Wireless Access
TCP = Transmission Control Protocol
UDP = User Datagram Protocol
VoIP = Voice over IP
WAP = Wireless Application Protocol

2 Traffic Modeling for MBWA Simulations

2.1 Introduction

The Mobile Broadband Wireless Access (MBWA) systems being discussed in IEEE 802.20 standards group are designed to provide a broadband, IP-oriented connection to a wireless user that is comparable to wired broadband connections that are in use today. It is expected that there will be a mix of user applications, not unlike that of such wired systems. Further, the traffic charactersitics and system requirements of the various applications can vary widely. The performance of such MBWA systems is thus very much dependant on the details of the applications and their traffic models. This is in contrast to cellular wireless voice systems where the performance studies focused on physical and link layer performance with a relatively simple traffic generation model. The purpose of this document is to provide detailed statistical traffic models which can be used as an input to generate packets in a simulation study of a MBWA system. It will be a companion to the detailed statistical channel models which specify the wireless transmission channel impairments.

2.2 Context and Scope

2.2.1 User scenarios

There can be various different user scenarios for MBWA systems, some of which we cannot foresee at this time. For purposes of illustration, we include some candidate scenarios to frame the context of our work. [Editor's note: These descriptions need to be discussed]. In all cases, the MBWA modem can either be built-in or supplied through a card or a peripheral device.

a) Laptop user: The large and rich display capabilities can be expected to generate graphics-rich and multimedia-rich applications. In general, laptop users will provide the highest data volume demands due to the storage and battery capabilities of laptops. They can provide a full range of applications with perhaps less emphasis on voice and WAP applications. Except for special cases, they tend to be stationary during use.

b) PDA user: The display, battery, and storage capabilities are less than that of laptops, and so they are expected to have somewhat less traffic volume. They can be very portable. They are typically used for Web browsing, e-mail, synchronization, video, and voice applications.

c) Smartphone: These devices are very portable and very constrained display and storage capabilities. It is expected that they will be oriented towards voice, WAP, and light video.

2.2.2 Basis for Traffic Models

Most traffic modeling work is based on measurements of real traffic, which are analyzed to generate usable statistical descriptions. These are typically used in computer simulations, but can also be used to generate packet traffic for a real system under test. Since MBWA is a future service that is similar to some existing wired systems, a lot of the basis of this document is the traffic modeling work done for wired systems. These provide a reasonable and realistic description of the potential user. Our approach is to use statistical models that can be used to generate a stream of packets that need to be transmitted over the system.

We realize that characteristics of user applications keep changing. At best, one can develop a reasonable consensus model that is useful for bringing some uniformity in comparisons of systems. In particular, it is known that user traffic patterns change as the network performance changes. Traffic modeling work has attempted to adjust to this trend. For example, some of the traffic models such as Web and FTP try to capture the essence of the user applications by describing the amount of data work the user is trying to retrieve rather than specifying a packet stream.

We specifically do not use the trace-based approach where a real recorded stream of packets is played back for simulation. While traces can capture sophisticated details, such traces have details that are often very dependant on the system from which they were recorded, and do not provide flexibility for computer simulation work.

2.2.3 Traffic Mix

A MBWA system is expected to have mix of traffic types. There can be different types of usage scenarios (multi-service v. single-type), different types of devices (laptops v. PDAs), different levels of use (intense v.

light)., and different demands on response times (real-time v. best-effort). This document is primarily concerned with the traffic models for each of the potential traffic type. As discussed in the previous section, these are based on statistical analysis of measured traffic to extract some invariant patterns that are not very dependant on the specific system. It is more difficult to describe a similar invariant mix of traffic types since these tend to depend more heavily on the type of system and the mix of device/user types.

In the context of a system evaluation using traffic models, the specific mix of traffic types will emphasize different aspects of the system performance, e.g. sustained throughput for file downloads v. faster response times for interactive applications. [Editor's note: This needs to be discussed] While we may discuss some candidate traffic mixes as pre-mixed recipes for consideration, the recommended mix of traffic types may be outside the scope of this document since it may be intimately connected to the weightage given to various requirements of the MBWA system.

Ref [1, 2, 4] discuss traffic mixes. Refs [2, 4] have graphs of measured traffic mixes. Ref [1] C802.20-03/43 proposes a traffic mix with HTTP/FTP/WAP/real-time video, which is given here for illustration:

Table 2-1 Assignment Probability for Different Traffic Models

HTTP	FTP	WAP	Real Time Video
24.43%	9.29%	56.43%	9.85%

2.2.4 Adaptive applications

[Editor's note: This needs to be discussed]. Certain applications such as audio streaming sense the available bit rate of the channel and then adjust the amount of traffic that is transmitted. Certain multi-media sessions may employ content-adaptation of images or video based on network conditions. This directly changes the amount of data that is transmitted. The adaptive nature of applications can be incorporated into the traffic model.

2.2.5 Higher-layer protocols

The specific details regarding the use of higher layer protocols such as HTTP, TCP, UDP are outside the scope of this document. It is true that there are certain relationships between protocols and applications such as: Web-browsing/HTTP and FTP typically use TCP and Audio Streaming typically uses UDP. But there are various flavors of TCP, which we are not specifying. There are also some dependancies, e.g. HTTP v1.1 and v1.0 will create different types of packet streams for the same Web browsing model. However, we can still describe traffic models effectively, and leave the protocol decisions out of the scope of this document.

2.2.6 Performance requirements

The performance requirements for the applications being described here are beyond the scope of this document.

3 Traffic Models for MBWA

This section described the traffic models in detail. Sections 3.1 and 3.2 clarify some aspects of the modeling approach and the remaining sections provide detailed models for traffic type (see ref [3] for a candidate list of applications).

[Editor's note: There has been very little detailed traffic model contributions. We need to discuss the available models. There are also new traffic types where we need input from people. There are a few references at the end which may be useful for some new application types. Refs are all available on drop-box or public Web or through IEEE Explore. It might be useful to circulate useful references even if finished models are not available]

3.1 User/Traffic Modeling Approach

[Editor's note: Notion of "modeled" user and relation to active/idle/registered/non-registered user needs to be clarified. See Ref[6] for a definition of active user. Ref [2] outlines concepts of active, hold, sleep states.] One of the objectives of a modeling and simulation exercise is to determine the number of users a MBWA system can support. The proposed approach here is to have traffic models for a user who is maintaining a session with transmission activity. These can be used to determine the number of such registered users that can be supported. This document does not address the arrival process of such registered users, i.e. it does not address the statistics of subscribers that register and become active.

3.2 Packet Generation

In some of the traffic models, there is a statistical description of the workload or the content of the application rather than the actual packet stream. This is consistent with the state of the art in evaluation of multi-service data systems. For example, the Web browsing model describes the Web pages and the timing between the Web pages. Depending on the details of the underlying TCP model (e.g. MTU size, max receive window) and the HTTP (HTTP v1.0 v. HTTPv1.1), the actual stream of packets will change. In some cases, as in the Voice models, the model may describe the packet stream more directly.

3.3 Web Browsing

Web browsing is the dominant application for broadband data systems, and has been studied extensively. See references [1, 2, 7, 8].

(C802.20-03.43 (ref [1]) has a detailed description of a Web model which is based on ref [7]. The basic model is included here for illustration. See also ref[8].)

The parameters for the web browsing traffic are as follows:

S_M: Size of the main object in a page

S_E: Size of an embedded object in a page

N_d: Number of embedded objects in a page

D_{pc}: Reading time

T_p: Parsing time for the main page

Table 3-1 HTTP Traffic Model Parameters

Component	**Distribution**	**Parameters**	**PDF**

Main object size (S_M)	Truncated Lognormal	Mean = 10710 bytes Std. dev. = 25032 bytes Minimum = 100 bytes Maximum = 2 Mbytes	$f_x = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left[\frac{-(\ln x-\mu)^2}{2\sigma^2}\right], x \geq 0$ $\sigma = 1.37, \mu = 8.35$
Embedded object size (S_E)	Truncated Lognormal	Mean = 7758 bytes Std. dev. = 126168 bytes Minimum = 50 bytes Maximum = 2 Mbytes	$f_x = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left[\frac{-(\ln x-\mu)^2}{2\sigma^2}\right], x \geq 0$ $\sigma = 2.36, \mu = 6.17$
Number of embedded objects per page (N_d)	Truncated Pareto	Mean = 5.64 Max. = 53	Note: Subtract k from the generated r.v. $f_x = \frac{\alpha k^{\alpha}}{x^{\alpha+1}}, k \leq x < m$ $f_x = \left(\frac{k}{m}\right)^{\alpha}, x = m$ $\alpha = 1.1, k = 2, m = 55$ to get N_d
Reading time (D_{pc})	Exponential	Mean = 30 sec	$f_x = \lambda e^{-\lambda x}, x \geq 0$ $\lambda = 0.033$
Parsing time (T_p)	Exponential	Mean = 0.13 sec	$f_x = \lambda e^{-\lambda x}, x \geq 0$ $\lambda = 7.69$

Note: When generating a random sample from a truncated distribution, discard the random sample when it is outside the valid interval and regenerate another random sample.

3.4 FTP

(C802.20-03.43 (ref [1]) has a detailed description of an FTP model which is based on ref [7]. The basic model is included here for illustration. Also see ref [8])

In FTP applications, a session consists of a sequence of file transfers, separated by *reading times*. The two main parameters of an FTP session are:

S : the size of a file to be transferred

D_{pc} : reading time, i.e., the time interval between end of download of the previous file and the user request for the next file.

The underlying transport protocol for FTP is TCP. The parameters for the FTP application session are described in Table 3-2.

Table 3-2 FTP Traffic Model Parameters

Component	**Distribution**	**Parameters**	**PDF**
File size (S)	Truncated Lognormal	Mean = 2Mbytes Std. Dev. = 0.722 Mbytes Maximum = 5 Mbytes	$f_x = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[\frac{-(\ln x - \mu)^2}{2\sigma^2}\right], x \ge 0$ $\sigma = 0.35, \mu = 14.45$
Reading time (D_{pc})	Exponential	Mean = 180 sec.	$f_x = \lambda e^{-\lambda x}, x \ge 0$ $\lambda = 0.006$

Based on the results on packet size distribution 76% of the files are transferred using and MTU of 1500 bytes and 24% of the files are transferred using an MTU of 576 bytes. For each file transfer a new TCP connection is used whose initial congestion window size is 1 segment (i.e. MTU). The packet arrival process at the base station is described by the TCP model described earlier. The process for generation of FTP traffic is described Figure 1.

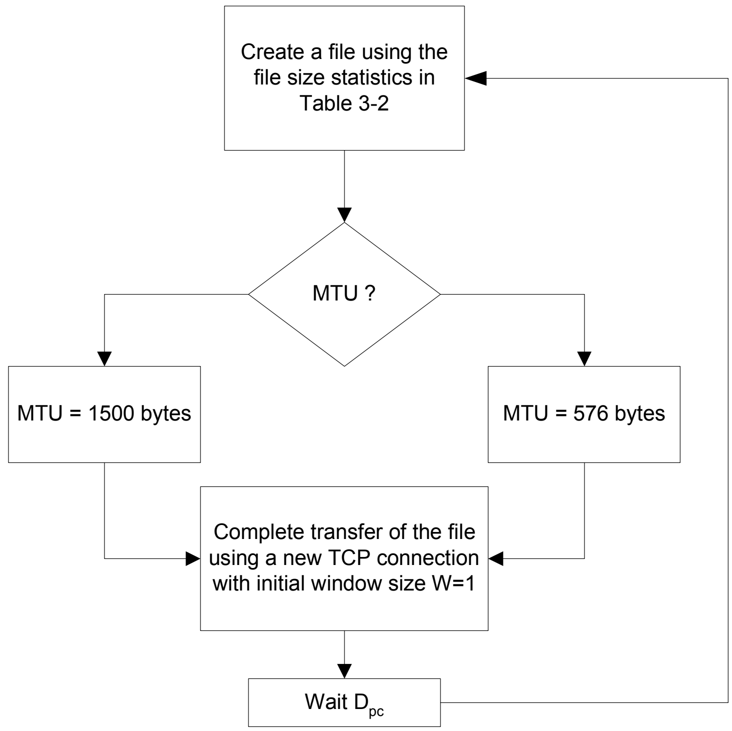

Figure 1 Model for generating FTP traffic

3.5 E-mail

[Note: E-mail is an important application for any Internet access system. See ref [8]]

3.6 WAP

[Following content is pasted from C802.20-03.43 for illustration]

Each WAP request from the browser is modeled as having a fixed size and causes the WAP server to send back a response with an exponentially distributed response time. The WAP gateway response time is the time between when the last octet of the request is sent and when the first octet of the response is received from the WAP server. The response itself is composed of a geometrically distributed number of objects, and the inter-arrival time between these objects is exponentially distributed. Once the last object is received, the exponentially distributed reading time starts, and it ends when the WAP browser generates the next request. Table 3-3 describes the distribution of the model parameters. During the simulation period, the model assumes that each WAP user is continuously active, i.e., making WAP requests, waiting for the response, waiting the reading time, and then making the next request.

Table 3-3 WAP Traffic Model Parameters

Packet based information types	Size of WAP request	Object size	# of objects per response	Inter-arrival time between objects	WAP gateway response time	Reading time
Distribution	Deterministic	Truncated Pareto	Geometric plus offset of 1	Exponential	Exponential	Exponential

		(Mean= 256 bytes, Max= 1400 bytes)				
Distribution Parameters	76 octets	K = 71.7 bytes, α = 1.1	Mean = 2 plus offset of 1	Mean = 1.6 s	Mean = 2.5 s	Mean = 5.5 s

3.7 Voice (VoIP)

[Editor's note: Has the following assumption been made?] Since the intent of MBWA is to be an IP-oriented access system, voice applications are likely to be VoIP-based.

[Following from C802.20-03.43 is included for illustration. See [7]]

The voice traffic model will depend on the voice codec used as well as whether voice in 802.20 will be implemented as a circuit switched or voice over IP. Voice will in general follow a Markov source model with different rates (full rate, half rate, etc) with a corresponding set of transition probabilities between different rates.

3.8 Video (Videotelephony/Videoconferencing)

3.9 Audio streaming

This can be an important class of traffic. It has received relatively less attention in the modeling community. (See ref [9])

3.10 Video streaming

(C802.20-03.43 (ref [1]) has a detailed description of a video streaming model which is based on ref [7]. The basic model is included here for illustration.)

The following section describes a model for streaming video traffic on the forward link. Figure 2 describes the steady state of video streaming traffic from the network as seen by the base station. Latency of starting up the call is not considered in this steady state model.

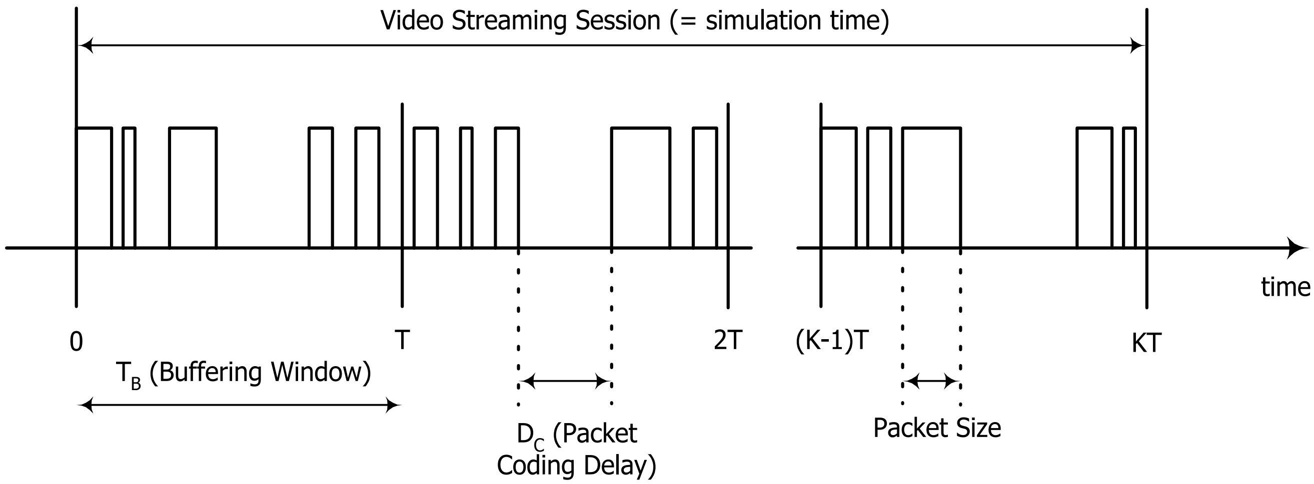

Figure 2 Near Real-Time Video Traffic Model

A video streaming session is defined as the entire video streaming call time, which is equal to the simulation time for this model.

Each frame of video data arrives at a regular interval T determined by the number of frames per second (fps). Each frame is decomposed into a fixed number of slices, each transmitted as a single packet. The size of these packets/slices is distributed as a truncated Pareto. Encoding delay, Dc, at the video encoder introduces delay intervals between the packets of a frame. These intervals are modeled by a truncated Pareto distribution. The parameter T_B is the length (in seconds) of the de-jitter buffer window in the mobile station used to guarantee a continuous display of video streaming data. This parameter is not relevant for generating the traffic distribution but is useful for identifying periods when the real-time constraint of this service is not met. At the beginning of the simulation, it is assumed that the mobile station de-jitter buffer is full with (T_B x source video data rate) bits of data. Over the simulation time, data is "leaked" out of this buffer at the source video data rate and "filled" as forward link traffic reaches the mobile station. As a performance criterion, the mobile station can record the length of time, if any, during which the de-jitter buffer runs dry. The de-jitter buffer window for the video streaming service is 5 seconds.

Using a source video rate of 32 kbps, the video traffic model parameters are defined Table 3-4

Table 3-4 Near Real-Time Video Traffic Model Parameters

Information types	Inter-arrival time between the beginning of each frame	Number of packets (slices) in a frame	Packet (slice) size	Inter-arrival time between packets (slices) in a frame
Distribution	Deterministic (Based on 10fps)	Deterministic	Truncated Pareto (Mean= 50bytes, Max= 125bytes)	Truncated Pareto (Mean= 6ms, Max= 12.5ms)
Distribution Parameters	100ms	8	K = 20bytes $\alpha = 1.2$	K = 2.5ms $\alpha = 1.2$

3.11 Gaming

Some types of multi-player games may have demanding requirements on response times. This traffic type has been mentioned in ref [3].

(Editor's note: See refs [7, 10, 11])

3.12 Other traffic types

(Editor's note: To be discussed and added: e.g. File-sharing (P2P), PDA synchronization, NNTP ??)

4 References

[1] C802.20-03/43, "802.20 Evaluation Methodology Strawman (03/57 is ppt)", IEEE 802.20 May 2003 Session.

[2] C802.20-03/13r1, "User Data Models for an IP-based Cellular Network," IEEE 802.20 March 2003 Session.

[3] C802.20-03/35, "Evaluation Methodology for MBWA", IEEE 802.20 May 2003 Session.

[4] C802.20-03/53, "Operator Systems Requirements for MBWA", IEEE 802.20 May 2003 Session.

[5] C802.20-03/46r1, "Channel Requirements For MBWA (Rev 1)S", IEEE 802.20 May 2003 Session.

[6] IEEE 802.20 Requirements Document v2.

[7] 1xEV-DV Evaluatuion Methodogy, 3GPP2/TSG-C.R1002. Document available in 802.20 drop-box.

[8] D. Staehle et al, "Source Traffic Modeling of Wireless Applications," Research Report, June 2000. available at: http://www3.informatik.uni-wuerzburg.de/TR/tr261.pdf

[9] A. Mena and J. Heidemann, "An empirical study of real audio traffic," INFOCOM 2000. Proceedings. IEEE , volume: 1 , 26-30 March 2000. pp 101 -110 vol.1

[10] R. A. Bangun and E. Dutkiewicz, "Modelling multi-player games traffic," Information Technology: Coding and Computing, 2000. Proceedings. International Conference on , 27-29 March 2000. pp: 228 -233

[11] C. Heyaime-Duverge and V. K. Prabhu, "Modeling action and strategy Internet-games traffic Vehicular Technology Conference, 2002. VTC Spring 2002., Vol. 3 , 6-9 May 2002, pp: 1405 -1409 vol.3