IEEE P802.11

WIRELESS LAN

DATE: **NOVEMBER 13, 1991**

TITLE: **COMMUNICATIONS REQUIREMENTS OF MULTIMEDIA APPLICATIONS: A PRELIMINARY STUDY**

AUTHOR: Timothy Kwok
APPLE COMPUTER, INC.
20450 Stevens Creek Blvd., MS: 76-2H
Cupertino, CA 95014
(408) 974-8311
Fax (408) 974-9793

ABSTRACT

A preliminary study is presented that provides a framework to explore the communications requirements of wireless applications in general, with emphasis on multimedia applications. An estimate of bandwidth requirements for various basic common applications are given, from which network bandwidth requirements in different scenario can be evaluated. The delay and error characteristics for certain multimedia applications are described, along with possible implications about the suitability of random access and reservation-based MAC protocols.

COMMUNICATIONS REQUIREMENTS OF MULTIMEDIA APPLICATIONS: A PRELIMINARY STUDY

Timothy Kwok
Apple Computer, Inc.

ABSTRACT

A preliminary study is presented that provides a framework to explore the communications requirements of wireless applications in general, with emphasis on multimedia applications. An estimate of bandwidth requirements for various basic common applications are given, from which network bandwidth requirements in different scenario can be evaluated. The delay and error characteristics for certain multimedia applications are described, along with possible implications about the suitability of random access and reservation-based MAC protocols.

I. INTRODUCTION

Communications networks are designed for the needs of providing access to people, information and processing power [1] that are remotely located, as well as sharing expensive and inherently limited resource such as interconnection bandwidth ("spectrum") and printing devices. Such needs are not mutually exclusive; for example, access to supercomputing power by various users over the network is also a means of sharing the expensive supercomputing resource. The ultimate, utopian goal of telecommunications has been to set the human free from the constraints of distance, location (mobility of person, entity, service, terminal), time and the medium of communication.

Distance. The issue of distance addresses the means of communication between individuals separated geographically. Since the invention of the telephone many decades ago, the worldwide wired telephone network has evolved to a nearly-ubiquitous presence. The constraint of distance is generally removed as far as voice communication as concerned.

Location. However, even the wired telephone infrastructure ties a person to a telephone at a fixed location. To make outgoing calls, one has to find a location with a telephone connection. To receive a telephone call, one has to be at the location where his phone number has been registered to the network, e.g., one's home or office. Only during the last decade has the constraint of location been addressed effectively, with the development of a wireless "cellular" telephony infrastructure. This allows a person to communicate wherever he is, independent of location (but at a premium price).

Time. During the same time frame, answering machine has removed us from the constraint of time: the recipient of a message does not have to be present at the time of at least the initial (one-way) communication.

Medium. Finally, the quality of communication depends on the medium of communication. Traditionally, communications are mainly in the form of voice. More recently, text communication in the form of electronic mail and low quality image and video communication in the form of facsimile and video conferencing have also been introduced, utilizing the telephony infrastructure. However, these functions have been added to the voice-optimized network in a very clumsy manner. The ideal medium of communication should allow unconstrained exchange of text, audio, image and video information to support multimedia and collaborative applications easily and through a user friendly interface.

An enormous international effort has recently emerged to create a worldwide information infrastructure referred to as B-ISDN [2], based on a standardized set of networking protocols, to support a very wide range of applications. This development is expected to guide the future of computing and communications in the 1990s and beyond and can be particularly effective for supporting multimedia and collaborative applications.

Simultaneously, another development that has received significant attention, both in the U.S. and worldwide, is personal communications service (PCS). Initially, PCS was narrowly described as an evolution from the current analog cellular telephone service (by using microcellular digital technologies and lower power mobile phones) to allow more affordable provision of mobile services to the general public. Refined visions of PCS have been proposed in [3] to allow a person to direct calls to either another specified person via a personal number or a specified place via a place number.

While PCS has the emphasis of voice telephony, Data PCS as proposed by Apple [4] emphasizes the importance of providing data (which includes all information types, e.g., voice is a type of data) connectivity. Functionally and at its mature development phase, Data-PCS can be envisioned as providing wireless capabilities similar to those provided on cables by B-ISDN: i.e., a wireless infrastructure capable of supporting a very wide range of applications.

Some parties have suggested that as such, Data-PCS may well share similar technology employed by B-ISDN, for example, ATM (asynchronous transfer mode) as the transport mechanism for all applications. However, there are two fundamental attributes in the wireless environment that are different from a wired network: a finite supply (and thus scarcity) of bandwidth, and unreliable transport. ATM technologies have been developed assuming operation in the fiber optics environment, for which bandwidth is abundant and transport is reliable. To use the sophisticated protocols of ATM in a wireless environment, significant modifications in the ATM protocol may be needed to account for the changed assumptions. The topic, however, merits further study.

Designing a universal network such as B-ISDN and Data-PCS to support all envisioned applications requires an in-depth understanding of those applications. The objective of this paper is to suggest a means to analyze and characterize applications so that networking schemes can be designed to effectively support them, both in the wired and wireless environment. In section II of this paper, we will define what an application is. In section III, a communication model is presented to describe the assumptions of the following discussion. Then, section IV will explain how an application in general can be characterized based on its traffic generation pattern and its corresponding communication requirements. A classification scheme for applications based on their communications requirements is presented in section V. Finally, section VI discuss the implications on supporting multimedia applications on MAC protocol design.

II. WHAT IS AN APPLICATION ?

An application is defined as a task that requires communication of one or more information streams between two or more parties that are geographically separated.[1] More specifically, an application can be characterized by the following attributes (see Table I):

(1) INFORMATION TYPES

In general, information can be classified as time-based and non-time-based. Time-based information is defined as those that must be presented at specific instants to convey its meaning. Typical time-based information are video and audio, while non-time-based information includes still images,

1 Here we focus on communications applications only, and not standalone applications.

graphics, text, etc. An application can in general include both time-based and non-time-based information. Synchronization is an important issue when an application involves simultaneous transfer of different information sources [5].

(2) DELIVERY REQUIREMENT

Applications can be classified according to its delivery requirements into real-time and non-real-time applications. A real-time application is defined as one that involves information delivery for immediate consumption. In contrast, for non-real-time applications, the information is stored at the receiving party for later consumption. The former requires sufficient bandwidth, while the latter requires sufficient storage. (Conversation on the telephone network is considered real-time applications, while leaving a voice mail message on an answering machine is a non-real-time application.) In other words, communicating parties for a real-time application participate at the same time, while for a non-real-time application participate at different time.

(3) SYMMETRY OF CONNECTION

In general, a communication application involves a two-way transfer between a communicating parties. Such bidirectional connection can be classified as either symmetric and asymmetric connection. A symmetric connection is one that involves transfer of information of similar traffic characteristics of both directions, otherwise, it is called an asymmetric connection. (A voice call is a good example of a symmetric connection, while video browsing is considered as an asymmetric connection that involves sending control messages in one direction and video transfer in the other direction. Video broadcast is an extreme case of asymmetric connection that involves a one way communication.)

(4) POINT-TO-POINT VS MULTIPOINT

Applications can be classified into point-to-point or multipoint depending on the number of parties involved in the communication; those involving two parties are called point-to-point, while those involving more than two parties are called multipoint. Traditional phone conversations are point-to-point applications, whereas teleconferencing that involves three or more parties are multipoint applications. Broadcast is the extreme case of multipoint application that involves sending information to all parties of t network. One of the main issues in multipoint connection is the brigding function, namely, the distribution of information streams to more than one destination. Such function can be performed by either a dedicated point within the network, or provided by one of the communicating parties. Both approach have different implications on network bandwidth requirement and reliability.

(5) HUMAN VS COMPUTING DEVICE

In general, the parties involved in the application may be either a human user (through a user terminal) or a computing device. Again, a voice call is a user-user application, while a user accessing a remote database is a user-devi application, and finally two supercomputing performing parallel computation to solve the same problems communicating intermediate results is considered as a device-device application. Collaboration [6] is an application between human users that shares an electronic space for communication, which can be either real-time or non-real-time.

(6) ENVIRONMENT

The network that connect the geographically separated parties can be either wired or wireless. The need for wireless networks stems from the need of communication independent of the location of the communicating parties; the parties are not limited to the fixed locations where wired networks terminate. (An obvious examples of wireless networks is the cellular telephone networks, while the public switched telephone network is a wired network.)

(7) MOBILITY

Most communication take place between fixed locations and individual are tied to the predetermined location to obtain the communication service. Mobility can be described as the ability to use a single logical label to access an entity wherever it is within the service area of the network. Mobility is a concept that can be applied to a person, terminal and service [7]. Personal mobility allows connection to a person through the network independent of the location of that person. Terminal mobility allows the same capability for the physical terminal. In this case, more than one person can be assigned to same the same terminal. Service mobility allows communication to a service, which is provided at multiple locations, that is closest to the caller. In general, mobility requires an intelligence within the network to know the location of the entity, such as by keeping a database of the current physical location of the entity for translation. Although mobility is a concept that is usually associated with wireless networks, mobility does exist today in the PSTN, such as the 911 number service, where calls are automatically routed to the closest emergency service center to the calling party.

ATTRIBUTES	CLASSES	
INFORMATION TYPES	**Time-based**	**Non-time-based**
	video, voice	images, text
DELIVERY REQUIREMENT	**Real-time**	**Non-real-time**
	telephone call	downloading video
CONNECTION	**Symmetric**	**Asymmetric**
	telephone calls	Video-on demand
NUMBER OF PARTIES	**Point-to-point**	**Multipoint**
	2-party call	conference call
COMMUNICAT-ING PARTIES	**Human**	**Device**
	Student	Database
ENVIRONMENT	**Wired**	**Wireless**
	Phone network	Cellular Telephony
MOBILITY	**Fixed**	**Mobile**
	POTS	Cellular phone

Attributes of an application and examples.
TABLE I

Relationship between Information Types and Delivery Requirements

As mentioned before, applications can be classified broadly into two types of delivery requirements: real-time and non-real-time. Real-time applications have stringent delay requirements for communication thus requiring sufficient bandwidth, while non-real-time applications are flexible in terms of delay because information will be stored for later use. Video conferencing and image browsing are typical examples of real-time applications, while downloading digitized movies and electronic mail belong to non-real-time applications. However, it is important to distinguish between the delay requirement of the application from the intrinsic time dependency of its information content, which can be time-based or non-time-based. The examples given for real-time and non-real-time applications have been chosen to include both time-based and non-time-based information types (see Table II). In the case of image browsing for real-time applications, even though the image is itself non-time-based information, to ensure interactive response for the user, a maximum response time constraint is required to satisfy this application. On the other hand, for non-real-time applications like downloading a digitized movie, even though the information content is a time-based, the entire movie can be treated as a single file transfer like electronic mail because the movie is not being displayed in real-time at the receiver. Therefore, in terms of specifying delay requirement for an application, the information content of the application is not the only determining factor; the nature of the application (real-time or non-real-time) is a critical factor to determine the delay requirement.

Delay requirements of applications will be discussed in more detail in the following section on communication requirements.

DELIVERY REQUIREMENTS	INFORMATION TYPES	
	Time-based	**Nontime-based**
Real-time	Video conferencing	Distributive Computing Collaborative Computing Interactive browsing
Nonreal-time	Downloading movie	Electronic mail

Examples of applications with various information types and delivery requirements.
TABLE II

III. A NETWORK MODEL

Before characterizing the applications, we need to define a model of the communication environment that supports those applications. The model comprises two or more computing devices (possibly includes, computers, storage, input and output devices, etc.) connected by a network. Each computing device may or may not have an associated human user. Since packet switching has emerged to be the preferred switching methodology compared to circuit switching because of the flexibility in supporting a wide variety of application with diverse requirements [8], the network we are going to discuss in the following will be packet switched network. Moreover, the network is assumed to use a reservation-based protocol for which network resource (bandwidth and buffer) can be negotiated between the network and the users in order to guarantee the quality of service (QOS) of an application. Typically such protocol is referred to as connection oriented (the ATM ,asynchronous transfer mode, protocol employed in B-ISDN is a example of such protocol).

IV. APPLICATION CHARACTERIZATION

An application can be characterized by its **traffic characteristics and communications requirements.** Each unit of the information generated by an application has its associated set of communication requirements. The network resource and protocol requirements can be derived from the communication requirements with respect to the traffic characteristics. In this paper, we will focus mainly on the traffic characteristics and communications requirements of applications; a more detail understanding of how they translate into specific network resource and protocol requirements is still under research.

(A) Traffic Characteristics

The traffic characteristics of an application can be specified by its traffic generation process. The traffic generation process is basically a sequence of messages generated at a arbitrary instants, each message having an arbitrary length. In the other words, it can be modelled as an on-off source. The pattern can be characterized by two probability distribution functions: (a) probability that a message is generated, given no message is currently generated (b) probability that the message generated is a particular length. Note that for aggregate traffic patterns, the bandwidth of the current data requirement is also a random variable, because it is a sum of the number of traffic generation sources.

Traffic patterns can be classified into two general categories according to their instants of message generation: **periodic traffic** and **sporadic traffic.**

If the instants of message generation occurs at regular intervals, it is called periodic traffic patterns, otherwise it is called sporadic traffic. The reason this classification is useful is that all time-based information types are periodic traffic (if generated at real-time), while non-time-based information types tends to be sporadic traffic. (For conventional 64 kbps PCM audio, samples of 8 bits are taken at 125 μs intervals. For uncompressed full motion NTSC video, a video frame a fixed amount information is generated at regular interval 1/30th sec or 30 frames/s. For compressed video such as those generated by the MPEG algorithm [9], a video frame is still created at regular intervals of 1/30th sec for NTSC and 1/25th sec for PAL formats, except the amount of information generated is variable at each instant. The variability of the message length is a key issue in supporting compressed video over packet switched networks. Statistical gain can be taken advantaged, but care has to be taken not to introduce significant delay.

The invention of packet switching in the 1960's was motivated by a very important class of traffic patterns typical of data communications, which is often referred to as bursty traffic. Bursty traffic is characterized by a sequence of periods of random time interval of information generation separated by gaps of silence of random duration. Conventional data communications are bursty because they are typically file transfer, remote login or more recently, traffic generated by diskless workstations. In general, bursty traffic is also typical for applications transferring non-time-based information or non-real-time applications. The unpredictability of bursty traffic is the main culprit that needs to be dealt with in designing the packet switched network. However, the statistical gain made possible by multiplexing of the bursty traffic of traditional data communications through the network medium can make packet switching more efficient than circuit switching.

(B) Communications Requirements

To completely characterize an application, not only is its traffic generation process needs to be specified, but also are its communications requirements. These requirements fall into three categories: bandwidth, delay, and error. **Bandwidth is the dominant requirement: if bandwidth is insufficient, delay and error may be introduced, which then must be addressed by the network to assure the quality of service for the application.** (Error, it should be noted, can be introduced independent of the adequacy of bandwidth.)

(1) Bandwidth

In general, the issue of bandwidth (or throughput, capacity, etc.) arises when a resource is required to perform a certain task. In the case of communication application, the networking bandwidth consists of the combination of multiple networking resources. These resources include transmission, switching and processing bandwidth within and at the end points of the network (such as interface and data base servers). Networks exist not only for the sake of communication among all connected parties, but also for sharing expensive resources like transmission and switching facilities, as well as the computing, printing and storage devices at the end points of the network. Therefore, multiple access of all these resources requires efficient sharing mechanisms. For transmission and switching, such mechanisms are known as multiplexing and switching methodologies, respectively. Here, we concentrate the discussion on transmission bandwidth issue, assuming other types of bandwidth is not a bottleneck.

The information rate of an applications can be calculated directly from the traffic characteristics by measuring the amount of information generated per unit of time[2]. Generally speaking, the information rate of an application can be either constant bit rate or variable bit rate. (Constant bit rate applications include traditional PCM coding of voice that generates 64kbps, as well as current video conferencing systems at 56 kbps, 2x56 kbps or p x 64 kbps. Conventional data communications almost always considered to be bursty in nature.) However, a lot of the traditional constant bit rate applications are intrinsically variable bit rate. Even voice itself can be treated as being variable bit rate in nature because voice can be sampled only when someone is talking and nothing needs to be sent otherwise. Similarly, video information needs to be sent only for the changes in image content, thus video compression algorithms intrinsically generates variable bit rate traffic. However, most applications that are intrinsically variable bit rate are encoded as constant bit rates because of the current requirements of the network. Thus today, applications are designed to move across the existing networks, while a more ideal case should be networks designed to support applications, which have diverse requirements.

2 The duration of this unit time is application dependent.

For a real-time application that communicates time-based information, the bandwidth required at each networking resource is characterized by its information rate. The bandwidth required may change as a function of time, like variable bit rate applications, or may be constant, like constant bit rate applications. If the bandwidth is insufficient, buffering will be required to avoid (random) delay —jitter, or if buffering is insufficient, then it may be necessary to drop information, a process which introduces error. This is one dimension where the bandwidth requirement is related to the delay and error requirements. If the real-time application is sending non-time-based information, the bandwidth requirement is a function of the response time requirement and the amount of information being communicated.

For non-real-time applications, the bandwidth requirement is simply a function of the response time and the amount of information that is communicating (the transmission time is usually insignificant relative to the response time required by non-real-time application).

Therefore, the network bandwidth required by each application depends on whether the delivery requirements as well as the information types (see Table III).

DELIVERY REQUIREMENT	INFORMATION TYPES	
	Time-based	**Non-time-based**
Real-time	Traffic Generation Rate	Response Time & Information Volume
Non-real-time	Response Time & Information Vol.	Response Time & Information Volume

Factors determining bandwidth requirements in point-to-point application.

TABLE III

In the (single hop) wireless environment, the communication network must almost always be considered as a shared medium. All the users on the network share access to that medium, and each message sent by any user is automatically broadcast to everybody.[3] This has two implications: an advantage and a disadvantage. First, the bandwidth resources needed in multicast or broadcast connection is the same as that for point-to-point connection, because every user automatically receive all messages sent out by any user (and discards those messages not addressed to itself). Second, the bandwidth cost of a two way connection is the sum of the two one-way connections, because each party of the two-way connection needs to consume the bandwidth from the same available medium. Therefore, the total bandwidth requirement on a wireless network depends on, in addition to the delivery requirement and information types of the application, the number of users on the network, the number of parties involved in each application (point-to-point vs multipoint) as well as the connection types of each application (i.e., whether it is symmetric or asymmetric).

To determine the network bandwidth required in each scenario, it is useful to find the bandwidth requirements of a set of basic point-to-point one-way applications that comprise the various information types and delivery requirements. The total network bandwidth requirement is equal to the sum of various combination of these basic application depending on the number of users as well as the different attributes of the applications. Such basic applications are shown in Table IV. Note that, however, the bandwidth requirement derived from response time is a peak bandwidth requirement, and the duration of such requirement is a function of the total amount of information to be transferred. The bandwidth requirements of each of these applications are shown in Table V.

DELIVERY REQUIREMENT	INFORMATION TYPES	
	Time-based	**Nontime-based**
Real-time	Video, Audio	Images, Graphics, Text, Commands
Nonreal-time	Video, Audio files	Images, Graphics, Text, Commands

Factors determining bandwidth requirements in point-to-point application.
TABLE IV

REAL-TIME DELIVERY OF TIME-BASED APPLICATIONS

Video	**Uncompressed**	**Compressed[4]**
videoconference	~ 200 Mbps	p x 64 kbps (H.261)[10]
NTSC	~ 200 Mbps	1.5 Mbps (MPEG)[9]
HDTV	~ 1 Gbps	20 Mbps [11,12]
Audio	**Uncompressed**	**Compressed**
voice telephony	64 kbps	16 kbps
CD Quality Stereo	1.4 Mbps (2x706 kbps)	256kbps (MPEG) (2x128kbps) [13]

NONREAL-TIME DELIVERY OF TIME-BASED APPLICATIONS

Video (2 hours)	Uncompressed movie size	Compressed movie size	Peak[5,6] Bandwidth (Response time = 15 min[7])
NTSC	180 Gbytes	1.35 Gbytes	12 Mbps
HDTV	900 Gbytes	18 Gbytes	160 Mbps

3 For this discussion, we assume there is no hidden node problem and there is no interference from neighboring wireless networks. These two issues will be addressed separately in later contributions.

4 An additional factor of compression can be achieved if a small screen size is needed.

5 This assume the local storage device requires the same bandwidth requirements.

6 Depends on the response time requirement.

7 Less than the typical time required to go the local video store to rent a video.

REAL-TIME AND NONREAL-TIME DELIVERY OF NONTIME-BASED APPLICATIONS

Images	Uncompressed Mbytes	Compressed (JPEG) Mbytes	Peak Bandwidth[5,6] Response time 0.1 sec	10 mins
Photo:[8] 1k x 1k x24 bit	3	0.06 - 0.3 (Lossy)	4.8 - 24 kbps	8 - 40 kbps
X-ray: 2kx2k x12 bit	6	3 (Lossless)	240Mbps[9]	0.4 Mbps

Bandwidth requirements in point-to-point one way application.
TABLE V

The total bandwidth requirement of each scenario can be determined by summing all the bandwidth required for each connection supported by the network. (Again, the bandwidth requirement for each connection depends on whether it is one way or two way connection, point-to-point or multipoint, and for multipoint, whether it is simple multicast or each point has different information content that require separate transmission.) Table VI shows the number of connections of various applications that can be supported by a 10 Mbps medium (assuming 100% network utilization is possible, i.e., these are upper bound estimates). For response-time-driven bandwidth requirements, the table shows the number of transactions (movies, images delivered) that can be supported by a network over a period of time, say one hour for video delivery and one minute for image delivery.

8 Actual photo size (resoluation) and color depth may varies.

9 This can be reduced by allowing response time to increase from the stringent 0.1 sec.

Applications (All compressed)	Number of concurrent connections
Voice conversation	3125
CD Quality Stereo	39
Videotelephony	
64 kbps	781
384 kbps	130
MPEG video one-way realtime delivery (point-to-point/multicast/broadcast)	6
	Number of transactions
Nonreal-time video delivery	**per hour**
MPEG video	3
HDTV	0.25
Image (compressed) delivery	**per minute**
1kx1kx24 photos	250 - 1250
2kx2kx12bit X-rays	25

Maximum number of concurrent connections or transaction rates supported on a 10 Mbps channel.
TABLE VI

(2) Delay

The issue of network delay mainly arises from insufficient instantaneous network bandwidth. If bandwidth is insufficient in any parts of the network, buffering is required (unless information is allowed to be dropped), a process which then introduced a random amount of the delay to the information being delivered.

For real-time delivery of time-based information, the delay requirements are absolute delay and permitted delay jitter. The absolute delay of a block of information within the information stream is defined as the time interval between the instant at which that the beginning of that block leaves the source, to the instant at which the beginning of that block arrives at the destination and is presented to the user. The absolute delay is important for real-time communication like video conferencing or conventional telephone conversation because feedback is expected within a certain time period for nature conversation to take place.

The delay jitter between two block of information is defined as the difference in time separation of the two blocks when they are generated at the sender from that when they arrive at the receiver. The delay jitter is important for time-based information because they must be presented at a certain rate for natural consumption by the user. The jitter constraint can be determined by the interval between consecutive samples of the time-based information when it is generated, because a sample is supposed to be displayed at the receiver after each of such interval. For example, 30 frames/sec video can allow a delay jitter of 33 msec, while tradition 8 kHz voice telephony can allow a 125 µsec delay jitter. However, if the absolute delay is large compared to delay jitter, for example, 1 sec absolute delay can be allowed for video, video frames can be buffered for that period to provide more flexibility for network delay jitter incurred, assuming there is sufficient cache at the receiver.

For non-time-based information or non-real-time delivery of time-based information, the major delay requirement is the absolute delay, which must be less than the response time required by the application.

(3) Error

There are two principal sources of error in the wireless environment: unreliability of the communication channels due to noise and unfavorable propagation conditions, and in the case of packet switching, buffer overflow at the either ends of the network (and at the switching nodes within the network if it is a multi-hop network). Packet switching, because of its statistical nature in multiplexing and switching, can introduce a random amount of the delay when the instantaneous bandwidth is not available at parts of the network and the information needs to the stored temporarily. Two sources of error may occur from this process. First, buffer may be insufficient and information needs to be discarded (retransmission may be needed depending on the information types). Second, in the case of time-based information, the delay introduced by buffering these information may exceed the delay jitter constraint, which will make that piece of information useless even when it finally arrived at the receiver (which makes retransmission futile); this is equivalent to the piece of information is dropped because of buffer overflow.

V. AN APPLICATIONS CLASSIFICATION

Since application can be characterized by its communications requirements in bandwidth, delay and error, a classification scheme for applications can be devised accordingly:

(A) Bandwidth: Stream vs Bursty

Applications can be classified, according to the type of bandwidth demands they place on the network, as either "stream" or "bursty applications". Stream type applications place a sustained bandwidth demand on a certain (fixed or variable) level for a continuous period of time, independent of the network bandwidth available. (Real-time delivery of time-based information is considered stream type applications.) Bursty type applications is characterized by large ratio of peak bandwidth to average bandwidth requirements.

(B) Delay: Sensitive vs Tolerant

Delay sensitive applications are defined as those that become useless if their information content are received later than a certain deadline. All real-time delivery applications can be considered as delay sensitive applications, because they have stringent delay requirements. These requirements can be examined from both macroscopic and microscopic viewpoints. Macroscopically, for non-real-time applications, delay is measured by latency involved in the transporting all the information to its destination. (For example, downloading movies from video database involve latency equal to the size of the movies divided by the bandwidth of the communication channel.) Typically, macroscopic delay requirements are satisfied at connection time by guaranteeing a minimum bandwidth for the connection. Microscopically, for real-time applications, delay is measured in terms of jitter. (For example, video conferencing requires frames of information delivered at the same rate they are generated. The delay of each frame of information can vary by at most a frame time, unless a smoothing buffer is introduced at the receiver end. Typically, if a frame is delayed by more than a frame time, such information will become useless for the receiver, which will skip such frame to display the following frame. To guarantee jittering within a frame time, bandwidth must be reserved equal to the peak bandwidth required by the network.)

(C) Error: Sensitive vs Tolerant

Similarly, error sensitive applications have tight error constraints. In fact, error sensitive information is defined as that require absolute integrity, i.e. zero error. In many cases, applications delivering non-time-based information are error-sensitive applications. If the network deliver faulty information, high level protocols must be applied to correct it. Absolute integrity is usually achieved by some combination of error correction coding or error detection with retransmission.

V. WIRELESS PROTOCOL REQUIREMENTS

Most multimedia applications involve real-time delivery of either time-based or non-time-based. To guarantee the QOS of such applications, instantaneous bandwidth must be available to minimize the delay that might otherwise be incurred. In other words, bandwidth needs to be reserved for such applications during the connection setup phase. Random-access based MAC protocol, such as ALOHA and CSMA types, may not of themselves be inherently capable of supporting those particular applications since they do not have built-in mechanisms for bandwidth reservation. (However, conceivably, in the case of the bandwidth requirement of the multimedia application is small compared to the aggregate network capacity, and the network is lightly loaded so that required instantaneous bandwidth is usually available, the QOS may still be acceptable.) One purpose of this paper is to facilitate the task of developing enhancements of random-access protocols to support a wider variety of multimedia, as well as to provide a framework for consideration of reservation-based MAC protocols.

VI. CONCLUSION AND FURTHER STUDIES

A preliminary study is presented that explores the communications requirements of applications in general, with emphasis on multimedia applications. Many of these applications are delay sensitive and require bandwidth reservation or other measures for guaranteed QOS. Random access MAC protocols may not be universally applicable for support of multimedia applications, but both random access and reservation-based MACs are fertile ground for further research. Ultimately, a complete mathematical description of how the traffic characteristics and associated communications requirements (bandwidth, delay and error) translate into network requirements (bandwidth and buffer) is required. This description includes the design of a good traffic descriptor set for negotiation between the terminal and network, as well as the monitoring of traffic once the application is supported. Only those studies that reflect such a more complete mathematical model will provide convincing evidence of the potential success of any specific MAC.

Finally, what have presented is only a framework, and many issues in the above discussion needs to be explored in more detail.

ACKNOWLEDGMENT

I would like to thank Jim Lovette for taking his precious time to review this paper and giving many good suggestions.

REFERENCES

[1] J. Patterson, C. Egido," Three Keys to the Broadband Future: A view of Applications," IEEE Network Magazine, March 1990, pp. 41-47.

[2] A. Day," International Standardization of BISDN," IEEE LTS Magazine, vol. 2, No. 3, Aug. 1991, pp. 13-20.

[3] D. Cox,"Personal Communications- A viewpoint," IEEE Communications Magazine, Nov 1990, pp. 8-20, 92.

[4] J. Lovette, "Data-PCS", Petition for Rulemaking to FCC for Amendment of Section 2.106, Jan 28, 1991.

[5] T. Little, A. Ghafoor, " Network Considerations for Distributed Multimedia Object Composition and Communication," IEEE Network Magazine, Nov 1990, pp. 32-49.

[6] M. Schrage," Shared Minds: The New Technologies of Collaboration," Random House, New York, 1990.

[7] R. Wolff, S. Parlamas, D. Hakim, M. Beller, " A Functional Model and Analysis of Personal Communications Services," to be published.

[8] S. Minzer, " Broadband ISDN and Asynchronous Transfer Mode (ATM)," IEEE Communications Magazine, Sept 1989, pp. 17-24,57.

[9] D. Le Gall, " MPEG: A Video Compression Standard for Multimedia Applications," Communications of the ACM, vol. 34, no. 4, April 1991, pp. 46-58.

[10] M. Liou, " Overview of the p x 64 kbit/s Video Coding Standard," Communications of the ACM, vol. 34, no. 4, April 1991, pp. 59-63.

[11] R. Jurgen," The challenges of digital HDTV," IEEE Spectrum Magazine, April 1991, pp. 28-30, 71-73.

[12] W. Luplow, " Digital High-Definition Television Takes Off," IEEE Spectrum Magazine, January 1991, pp. 65-68.

[13] H. Musmann," The ISO Audio Coding Standard," IEEE Globecom 1990, pp. 511-517.

[14] G. Wallace," The JPEG Still Picture Compression Standard," Communications of the ACM, vol. 34, no. 4, April 1991, pp. 30-44.