Proceedings of 5th Intl. Workshop on Mobile Multimedia Communications *MoMuc'98* October 12-14, 1998, Berlin

Achievable Dropping Rates under Variable Frame TDMA Schemes in the presence of Deadlines and Overhead.

César Santiváñez Ioannis Stavrakakis
Electrical and Computer Engineering Department
409 Dana Research Building, 360 Huntington Avenue
Northeastern University, Boston, MA 02115
e-mail: ioannis(cesar)@.cdsp.neu.edu

Abstract— **The objective of this paper is to determine the minimum system dropping rate induced by variable frame TDMA schemes supporting time-constrained applications with common maximum cell delay tolerance. Expressions are derived for the induced (optimal) system dropping rate and the maximum number of users that can be admitted in the network without violating the maximum dropping rate constrain is determined. In addition to the study of the optimal scheme, suboptimal TDMA schemes have been considered, such as a variable frame TDMA scheme in the presence of overhead and fixed frame TDMA schemes with and without overhead. The performance limiting factors associated with the suboptimal schemes are identified, and the magnitude of their (negative) impact is evaluated. Based on this information it is possible to point to performance improving modifications which should be pursued to the extent permitted by technological constraints.**

Keywords— **Multiaccess, Quality of service, TDMA, variable frame, overhead, deadline, dropping rate.**

I. Introduction

Wireless ATM is the natural choice for the local and broadband wireless network supporting services such as voice, data, image, and video [1], [2]. System architectures to enable "Wireless ATM" (WATM) have already been developed [3], [4], [5]. They employ a Data Link Control (DLC) layer to combat the unreliability of the wireless link, and a Medium Access Control (MAC) protocol to organize the sharing of the multiaccess channel.

MACs for WATM have been examined in [3], [6], [7], [8], [9], [10]. They all employ TDMA with on-demand assignment of the transmission resources by a central agent or *scheduler*. The scheduler needs to be communicated by the distributed users of their demands for resources, as well as inform the users of its decision (slot assignment). This communication takes place at discrete points in time (it is not continuous due to communication resource limitations) and is implemented by employing various mechanisms such as a control channel, piggy-backing on information bearing cells and polling procedures; such mechanisms introduce some overhead in the system.

For Wireless ATM networks, the scheduler has to allocate resources effectively by taking into consideration the QoS requirements of the supported applications. A call admission control function is assumed to be employed as well, shaping the set of the supported applications (accepted sessions). In this paper, the QoS is defined in terms of a maximum tolerable cell delay and dropping probability. A cell is dropped when its maximum tolerable delay is exceeded. Or, equivalently, when its remaining delay tolerance reaches zero and its service is not completed; the remaining delay tolerance is equal to the maximum delay tolerance when the cell is generated and decreases by one in every subsequent slot (cell service time unit). Such QoS parameters are typically associated with real-time applications such as voice and video.

The work presented in this paper is related to the MAC protocol developed under the ACTS Magic WAND project (MASCARA), [10]. The MASCARA protocol employs an adaptive TDMA scheme built around the concept of variable size frames. A frame is sub-divided in different periods, which are used for data transmission or exchange of critical information related to new arrivals and scheduler decisions. The performance of the MASCARA protocol has been evaluated by means of simulations which - although valuable - can not provide extensive insight into its behavior. In particular, it would be important to determine (analytically) the optimal achievable system dropping rate which would establish the performance limits of a variable frame TDMA scheme as well as its relative performance potential with respect to other related or alternative schemes.

The objective in this paper is to determine the minimum *system* dropping rate (or, equivalently, dropping probability) induced by variable frame TDMA schemes supporting time-constrained applications with common maximum cell delay tolerance. In addition, the performance of related suboptimal schemes (due to overhead or simplified (fixed) frame structure) is evaluated and the magnitude of the (negative) impact of the associated limiting factors is determined. Based on this study it is possible to point to performance improving modifications which should be pursued to the extent permitted by technological constraints.

The case of zero overhead is considered in Section II, where the Ideal Variable Frame Length TDMA (IVFL-TDMA) scheme is considered. The IVFL-TDMA scheme employs a frame structure of variable length, at the boundaries of which scheduling decisions are taken. A gated-STE service discipline is employed which starts servicing according to the STE (Shortest Time to Extinction [11]) service discipline the cells arrived before the beginning of the current frame only (gated), until all such cells are either served or dropped, marking the end of the current frame. That

This research is supported in part by a Fulbright Scholarship, the National Science Foundation, and the GTE Corporation

is, the scheduler services cells in the order of increasing remaining delay tolerance and drops expired (zero remaining delay tolerance) cells. The scheduler is assumed to have knowledge of the exact time when past arrivals occurred when scheduling decisions are taken only, as opposed to when these arrivals occur. Consequently, the gated-STE service discipline could schedule cells differently than the STE and, thus, suboptimally. This could be the case if an earlier arriving cell has a remaining delay tolerance greater than that of a later arriving cell at the time the later cell arrives. If such arrivals occur during different frames then the gated-STE service discipline would schedule for service the earlier cell first while the STE service discipline would schedule the later first. As a result, the performance of the gated-STE service discipline may be suboptimal. As long as always earlier arriving cells have a remaining delay tolerance less than or equal to that of a later arriving cell at the time the later cell arrives, then the scheduling decisions of the STE and gated-STE policies would coincide and the resulting *system* performance be identical. The above condition holds in the case in which all arrivals have the same maximum delay tolerance which is assumed to be the case in this paper.

In Section III, the more general case of nonzero overhead - referred to as the Real Variable Frame Length TDMA (RVFL-TDMA) scheme - is analyzed. RVFL-TDMA is similar to the IVFL-TDMA scheme, only that a fixed amount of time during each frame is utilized for request/assignment exchange between the users and the scheduler and not for cell transmission (frame overhead). In the present study the impact of the overhead is evaluated analytically by deriving tight bounds on the system dropping rate and a number of interesting comparative results are presented.

In Sections IV and V, two additional TDMA Schemes are included for comparison. In Section IV the Ideal Continuous Entry TDMA (ICE-TDMA) scheme is presented. The ICE-TDMA is equivalent to a (centralized) dynamic TDM as it would be employed in a fixed network node. No frame structure is present and the scheduler is assumed to have knowledge of all past arrivals at the time when they occur. Thus, cells are considered by the scheduler as soon as they are generated (Continuous Entry) which would not be feasible in a wireless environment (Ideal). It is assumed that the scheduler employs the STE service discipline. The closed form expression derived in [12] from the study of this scheme provides insight into the more realistic TDMA schemes in addition to determining a lower bound on the system dropping rate under any TDMA scheme supporting the same set of applications.

Section V presents the Real Fixed Frame Length TDMA (RFFL-TDMA) scheme. The RFFL-TDMA assumes a fixed frame length, a fixed part of which is considered to be overhead. It should be noticed that the RFFL-TDMA can be a non work conserving scheme since empty slots may be left in a frame while cells are waiting for service. Tight bounds on the induced system dropping rate - derived in [12] - are used to evaluate the (negative) impact of fixing

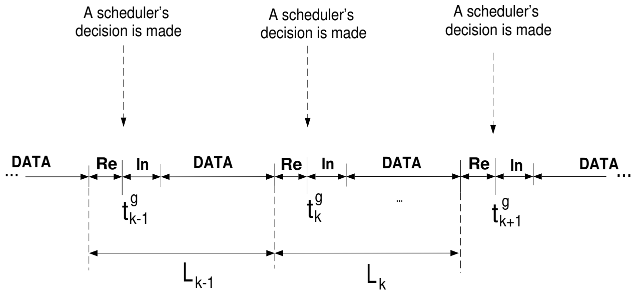

Fig. 1. Events and quantities of interest for the Real Variable Frame Length TDMA scheme.

the frame length. The study of the RFFL-TDMA scheme can be applied to the study of the Fixed Frame TDMA Scheme (FF-TDMA), where the scheduler's allocations do not vary on demand but they are static (quasi-static), as in T-1 systems (GSM sytems), eliminating (reducing) the need for reservations overhead. Once again, comparisons are presented between adaptive TDMA (RVFL-TDMA) and static TDMA (FF-TDMA).

The common environment assumed in all TDMA schemes consists of N memoryless time-constrained applications with identical maximum delay tolerance equal to T time units; the time unit is referred to as the (time) slot and is assumed to be equal to the cell service time. The N users compete for the time resource which is allocated by the scheduler at the scheduling decision times, as indicated earlier.

II. The Ideal Variable Frame Length TDMA (IVFL-TDMA) Scheme

The IVFL-TDMA scheme described briefly in the introduction is analyzed in this section.

Let t_k^g denote the instant of the *k-th* scheduling decision or beginning of the *k-th* service cycle (frame); let L_k denote the length of this frame ($L_k = t_{k+1}^g - t_k^g$) as shown in Figure 1 (with $Re = In = 0$). Without loss of generality, the arrival and departure processes are assumed to be right continuous. If no cell is waiting for transmission at time t_k^g, the present service cycle is empty and the next service cycle begins after one time slot. That is, an empty service cycle has a duration of 1 empty time slot. From the above and since any cell waiting for more than T time slots must be discarded, it follows that $1 \leq L_k \leq T$.

In the following subsections, the dropping rate is evaluated using the following procedure: First, the conditional expected number of dropped cells in a frame given that the length of the previous frame is L, namely $\bar{d}^I_{r/L}(T)$, is computed; superscript I stands for ideal. Second, using the transition probability matrix $< P^I_{ij} >$, describing the next frame length given the current one, the steady state frame length probability distribution Π^I_i is calculated. Finally, the dropping rate $d^I_r(T)$ is computed as follows:

$$d^I_r(T) = \frac{E_L\{\bar{d}^I_{r/L}(T)\}}{E\{L\}} = \frac{\sum_{i=1}^{T} \Pi^I_i \bar{d}^I_{r/i}(T)}{\sum_{i=1}^{T} i \Pi^I_i} \quad (1)$$

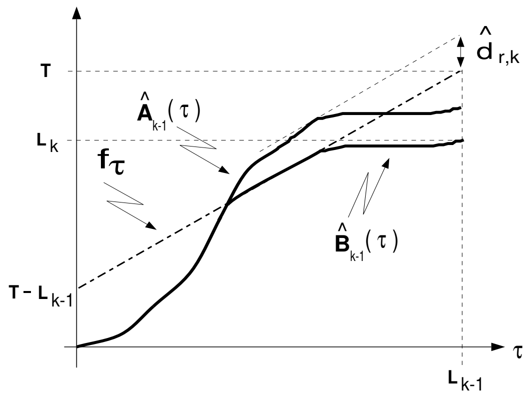

Fig. 2. An example of the evolution of $\hat{A}_{k-1}(\tau)$, $\hat{B}_{k-1}(\tau)$, and f_τ during the (k-1)-th service cycle

A. Conditional expected number of dropped cells, $d^I_{r/L}(T)$

Let $R(t)$ denote the cumulative arrivals up to (and including) time t (right continuous). Let $\hat{A}_{k-1}(\tau)$ (see Figure 2) denote the number of arrivals between time t^g_{k-1} and $t^g_{k-1}+\tau$, that is, τ time slots after the beginning of the previous service cycle:

$$\hat{A}_{k-1}(\tau) = \quad R(t_{k-1}+\tau) - R(t_{k-1}), \quad 0 \le \tau \le L_{k-1}.$$

Note that since sources are assumed to be memoryless, the evolution of the process $\{\hat{A}_{k-1}(\tau)\}_{\tau=0}^{L_{k-1}}$ does not depend on time t_{k-1} but only on the length of the previous cycle, L_{k-1}; also $\hat{A}_{k-1}(0) = 0$, since the arrival process is right continuous.

At time t^g_k the scheduler considers the $\hat{A}_{k-1}(L_{k-1})$ cells waiting for transmission, schedules some of them for transmission (namely L_k) and drops the rest (namely $\hat{d}_{r,k}$).

At time $t^g_{k-1}+1$ there will have arrived $\hat{A}_{k-1}(1)$ cells, the last of which will have to wait the $L_{k-1}-1$ time slots remaining for the end of the previous service cycle, plus $\hat{A}_{k-1}(1)$ time slots to complete its transmission. If this time $(\hat{A}_{k-1}(1) - 1 + L_{k-1})$ is greater than the maximum delay tolerance T, then $\hat{A}_{k-1}(1) - 1 - T + L_{k-1}$ cells will be dropped. In general, the number of cells that must be discarded by time $t^g_{k-1}+\tau$, namely $\hat{d}_k(\tau)$, is given by:

$$\hat{d}_k(\tau) \quad = \quad \max\{0, \max_{1\le\tau'\le\tau}[\hat{A}_{k-1}(\tau') - f_{\tau'}]\}$$

where $f_{\tau'} = \tau' + T - L_{k-1}, 0 \le \tau' \le L_{k-1}$. An example of the evolution of the process $\{\hat{A}_{k-1}(\tau)\}_{\tau=0}^{L_{k-1}}$ and the linear function f_τ is shown in Figure 2. Notice that f_τ represents the maximum number of arrivals up to time τ which can be transmitted before their deadline expires.

Let $\hat{m}_k(\tau) = \max_{1\le\tau'\le\tau}[\hat{A}_{k-1}(\tau') - \tau']$. Then $\hat{d}_k(\tau)$ can be rewritten as: $\hat{d}_k(\tau) = \max\{0, \hat{m}_k(\tau) - (T - L_{k-1})\}$. Notice that $\hat{d}_{r,k}$ - the total number of cells dropped at t_k, that is at the beginning of the *k-th* service cycle- will be equal to $\hat{d}_k(L_{k-1})$.

Let $A_j(\tau)$, $m(\tau)$, $d_j(\tau)$, and $d_{r,j}$ with $j = L_{k-1}$ be the random variables associated with $\hat{A}_{k-1}(\tau)$, $\hat{m}_k(\tau)$, $\hat{d}_k(\tau)$, and $\hat{d}_{r,k}$ respectively. These random variables do not depend of the time t^g_k but only on the previous cycle's length L_{k-1} described by j. $\{A_j(\tau)\}_{\tau=0}^{j}$ represents a right-continuous, discrete-valued, cumulative arrival process which has initial value zero ($A_j(0) = 0$), evolves for j time slots, and has independent increments. The probability of μ arrivals at any given discrete-time (increment) is given by l_μ.

Notice that $d_{r,j}$ denotes the number of cells dropped at the beginning of a service cycle that follows a service cycle of length j; its probability is given by:

$$P\{d_{r,j} = d\} \quad = \quad \begin{cases} P\{m(j) = T - j + d\} & \text{if } d \ge 1 \\ P\{m(j) \le T - j\} & \text{if } d = 0 \end{cases}$$

Let the function $P_j(\epsilon)$ be defined as $P_j(\epsilon) = P\{\max\{0, m(j)\} = \epsilon\}$. It is shown in [12] that this function can be computed using the recurrent formula:

$$P_j(\epsilon) \quad = \quad \begin{cases} \sum_{\mu=0}^{\min\{n,\epsilon+1\}} l_\mu P_{j-1}(\epsilon+1-\mu) & \text{if } \epsilon \ge 1 \\ l_0 P_{j-1}(0) + l_0 P_{j-1}(1) + l_1 P_{j-1}(0) & \text{if } \epsilon = 0 \\ 0 & \text{elsewhere} \end{cases}$$

with initial conditions: $P_0(\epsilon) = \delta(\epsilon)$ (i.e. zero always except at $P_0(0) = 1$).

Thus, the conditional expected number of dropped cells at the present frame given the previous frame length is L ($\bar{d}^I_{r/L}(T)$) is given by:

$$\bar{d}^I_{r/L}(T) = E\{\hat{d}_{r,k}/L_{k-1} = L\} = \sum_{d=1}^{+\infty} dP\{d_{r,L} = d\}$$
$$= \sum_{d=1}^{+\infty} dP\{m(L) = T - L + d\} = \sum_{d=1}^{+\infty} dP_L(T - L + d)$$

B. Service Cycle Length's Probability Distribution

The process $\{\hat{B}_{k-1}(\tau)\}_{\tau=0}^{L_{k-1}}$ with generic representation $\{B_j(\tau)\}_{\tau=0}^{j}$ (with $j = L_{k-1}$) will be considered. $\hat{B}_{k-1}(\tau)$ represents the number of 'surviving' (i.e. neither dropped nor assigned for transmission yet) cells among the cells arrived up to time $t^g_{k-1}+\tau$. Thus, for the variable frame length case:

$$\hat{B}_{k-1}(\tau) = \hat{A}_{k-1}(\tau) - \hat{d}_k(\tau) \quad \text{and} \quad B_j(\tau) = A_j(\tau) - d_j(\tau)$$

An example of the evolution of $\{\hat{B}_{k-1}(\tau)\}_{\tau=0}^{L_{k-1}}$ is shown in Figure 2. This evolution can be interpreted as if the cells that will have to be dropped, are being discarded as soon as this is realized, that is, each time $B_{k-1}(\tau)$ tends to become greater than the line $f_\tau = T - L_{k-1} + \tau$. It is clear that $L_k = \hat{B}_{k-1}(L_{k-1})$.
Let $PB^j_\tau(i,T) = P\{B_j(\tau) = i/$ maximum delay tolerance $= T\}$, then the following recurrent formula holds:

$$PB^j_\tau(i,T) = \begin{cases} \displaystyle\sum_{\mu=0}^{n} l_\mu PB^j_{\tau-1}(i-\mu,T) & \text{if } i < T - j + \tau \\ \displaystyle\sum_{\mu=1}^{n}(\sum_{\epsilon=\mu}^{n} l_\epsilon) PB^j_{\tau-1}(i-\mu,T) & \text{if } i = T - j + \tau \end{cases}$$

with initial condition $PB_0^j(i,T) = \delta(i)$ (1 at $i=0$, zero elsewhere). The first equation holds since no cell is dropped in the actual time slot (τ) when $i < T-j+\tau$; thus, $B_j(\tau) = B_j(\tau-1) + a_\tau$, where a_τ is the number of cells arriving at time τ (as before $l_\mu = P\{a_\tau = \mu\}$). The second equation holds since cells may be dropped when $i = T-j+\tau$. Thus, given that $B_j(\tau-1) = x$, then if (and only if) $a_\tau \geq T-j+\tau-x$ then $B_j(\tau) = T-j+\tau$.
From $PB_\tau^j(i,T)$, P_{ij}^I may be calculated as:

$$P_{ij}^I = \begin{cases} PB_j^j(i,T) & \text{if } i \geq 2 \\ PB_j^j(1,T) + PB_j^j(0,T) & \text{if } i = 1 \end{cases}$$

and then the steady state probability distribution of the service cycle length, $\Pi_i^I = P\{L_k = i\}$, can be computed.

Finally, the dropping rate $d_r^I(T)$ is computed from $\bar{d}_{r/L}^I(T)$ and Π_i^I by using equation (1).

III. The Real Variable Frame Length TDMA (RVFL-TDMA) Scheme

The RVFL-TDMA scheme is analyzed in this section (see Figure 1). The only difference between this scheme and the previous one (Section II) is the consideration of the *frame overhead*. In this real scheme, the first Re time slots at the beginning of every service cycle are consumed by the user's transmission requests; these Re time slots are referred to as the *reservation period*. The next In time slots are used by the scheduler to inform the users of its decisions (slot assignments) and are referred to as the *information period*. It is assumed that both Re and In are fixed and independent of the traffic load, but the following results can be easily extended to the case where the overhead period depends on the previous service cycle length.

The *k-th* service cycle begins at time $t_k^g - Re$, when the scheduler begins to receive the previous service cycle's arrival information of each user. At time t_k^g the scheduler has all the required information, takes its scheduling decision and informs the users during the next In time slots. At time $t_k^g + In$ the user's data transmissions (receptions) begin.

Let $t_{k,j}^g$ denote the time at which the *j-th* user completes the transmission of its request for slots of the $k-th$ frame, that is, provides to the scheduler information regarding its arrivals over the $(k-1)-th$ frame; $t_k^g - Re \leq t_{k,j}^g \leq t_k^g$. Clearly, this request will be based on user information (to be transmitted) which is generated before $t_{k,j}^g$. By assuming that this request (at $t_{k,j}^g$) represents all the information generated by user j until t_k^g or $t_k^g - Re$, the auxiliary systems L and U are constructed. That is, the following key assumptions are made regarding the content of the requests from all users, leading to systems L and U :

- *L :* At time t_k^g the scheduler knows about all arrivals up to time t_k^g.
- *U :* At time t_k^g the scheduler knows only about the arrivals up to time $t_k^g - Re$.

Under the auxiliary system L, a cell arriving over the interval $< t_{k,j}^g, t_k^g >$ will be considered for service during the $k-th$ service cycle. Under the real scheme this cell will be considered for service in the the $(k+1)-th$ service cycle. Clearly, the cell delay under the real scheme will be shaped by an additional service cycle length ($Re + In +$ data) compared to that under system L. Thus, the auxiliary system L outperforms the real system. A similar argument between the real system and auxiliary system U regarding cells generated over $< t_k^g - Re, t_{k,j}^g >$ establishes that the real scheme outperforms the auxiliary system U.

Let the superscripts I, R, L, U indicate a quantity associated with the IVFL-TDMA scheme, RVFL-TDMA scheme, auxiliary system L and auxiliary system U, respectively. In view of the above discussion it is easy to establish that: $d_r^I < d_r^L \leq d_r^R \leq d_r^U$.

The auxiliary systems L and U will be studied to calculate tight (as it will be shown) bounds on the performance induced by the real, variable frame length, gated scheme.

A. Computation of d_r^L

Since the maximum delay tolerance is T, the service of the last cell served over the $k-th$ service cycle must be completed by $t_k^g + T$. Thus, the maximum length of a busy (at least one served) frame will be equal to $t_k^g + T - (t_k^g - Re)$. Due to the overhead, the length of an empty frame will be equal to $Re + In$. Thus, $Re + In \leq L_k \leq Re + T$ for any k.

Let $\hat{A}_{k-1}(\tau)$, for $0 \leq \tau \leq L_{k-1}$ represents the number of cell arrivals (generations) over τ consecutive slots following t_{k-1}^g (as before); such arrivals will be considered for service during the $k-th$ frame (see Figure 1). Clearly, $\hat{A}_{k-1}(L_{k-1})$ represents all the arrivals (between t_{k-1}^g and t_k^g) to be considered for service during the $k-th$ frame.

Since $\hat{A}_{k-1}(\tau)$ cells have arrived at time $t_{k-1}^g + \tau$, then the last of these cells will have to wait $L_{k-1} - Re - \tau$ time slots (the remaining time for the end of the *(k-1)-th* service cycle) plus the overhead period ($Re + In$ time slots) plus $\hat{A}_{k-1}(\tau)$ time slots, in order to complete its transmission. Clearly, if this time ($\hat{A}_{k-1}(\tau) - \tau + L_{k-1} + In$) is greater that T time slots some cells will be dropped.
Thus, the number of cells arrived over $< t_{k-1}^g, t_{k-1}^g + \tau >$ which must be discarded is given by:

$$\begin{aligned}\hat{d}_k^L(\tau) &= \max\{0, \max_{1 \leq \tau' \leq \tau} [\hat{A}_{k-1}(\tau) - (\tau + T - In - L_{k-1})]\} \\ &= \max\{0, \max_{1 \leq \tau' \leq \tau} [\hat{A}_{k-1}(\tau) - (\tau + T' - L_{k-1})]\},\end{aligned}$$

where $T' = T - In$. For $L_{k-1} \leq T'$ this expression is the same to the one obtained in Section II.A. The results of Section II for the conditional dropping rate and the frame length are still applicable here. Two cases need to be considered :

<u>Case 1 :</u> $Re + In \leq j = L_{k-1} \leq T' = T - In$.
Processes $\hat{A}_{k-1}(\tau)$ and $\hat{B}_{k-1}(\tau)$ (defined in Section II) evolve as processes $A_j(\tau)$ and $B_j(\tau)$ with maximum delay tolerance $T' = T - In$, respectively. Taking into consideration that $L_k = \hat{B}_{k-1}(L_{k-1}) + Re + In$ the following "adjustments" to the quantities derived in Section II lead to the corresponding quantities associated with the auxiliary system L:

$$\bar{d}_{r/j}^L(T) = \bar{d}_{r/j}^I(T - I) \quad \text{, and}$$

$$P_{ij}^L \quad = \quad PB_j^j(i - Re - In, T - In)$$

where $\bar{d}_{r/j}^L(T)$ and P_{ij}^L denote the conditional expected number of dropped cells in the present frame given that the length of the previous frame is j, and the probability that the present frame length is i given that the previous frame length is j, respectively.

Case 2 : $T' = T - In < j = L_{k-1} \leq Re + T$.

The cells arriving at the first $j - T'$ time slots of frame $(k-1) - th$ will be discarded; for the remaining T' time slots, $\hat{A}_{k-1}(\tau')$ and $\hat{B}_{k-1}(\tau')$ will evolve as processes $A_j(\tau')$ and $B_j(\tau')$ with maximum delay tolerance T' and length $j' = T'$. Then:

$$\bar{d}_{r/j}^L(T) \quad = \quad (j - T + In)\lambda + \bar{d}_{r/T-In}^I(T - In) \qquad \text{, and}$$
$$P_{ij}^L \quad = \quad PB_{T-In}^{T-In}(i - Re - In, T - In)$$

where λ denotes the mean number of arrivals per slot.

Finally,

$$d_r^L(T) \quad = \quad \frac{\sum_{j=Re+In}^{Re+T} \Pi_j^L \bar{d}_{r/j}^L(T)}{\sum_{j=R+I}^{R+T} j\Pi_j^L} \qquad (2)$$

where $\Pi_i^L = P\{L_k = i \text{ at system L }\}$ and it is computed from P_{ij}^L.

B. Computation of d_r^U

In the auxiliary system L, arrivals over the overhead period $Re + In$ are treated differently. The arrivals over Re are served in the immediate service cycle while arrivals over In are served one service cycle later. In the auxiliary system U all arrivals over $Re + In$ are served later. Thus, the auxiliary system U can be viewed as equivalent to an auxiliary system L with parameters $Re' = 0$ and $In' = Re + In$. It is easy to establish that the performance of system U can be derived from the performance of the corresponding L system with delay tolerance reduced by Re. That is, $Re' = Re$, $In' = In$, and $T' = T - R$ and finally:

$$d_r^U(T) \quad = \quad d_r^L(T - Re)$$

IV. The Ideal Continuous Entry TDMA (ICE-TDMA) Scheme

Under the ICE-TDMA scheme the time-slots are (re)allocated every time a new cell arrival occurs (continuous-entry scheme). It is assumed that no time is consumed for the users to make their reservation or for the scheduler to inform the users of its assignments (ideal scheme). In [12] a closed form expression for the dropping as a function of the maximum delay tolerance T is derived. It turns out that the dropping rate can be described as the quotient between two IIR (Infinite Impulse Response) filters as follows (see [12]) :

$$d_r(T) \quad = \frac{[l_0 N(z)/D(z)]^{-1}}{[Y(z)]^{-1}}, \quad \text{where}$$

$$N(z) = \frac{z^{-1}[D(z) + \lambda - 1]}{1 - z^{-1}} \quad , \quad Y(z) = \frac{l_0 z^{-1}}{1 - z^{-1} D(z)}$$

$$D(z) \quad = \quad l_0 + (l_0 + l_1 - 1)z^{-1} + (l_0 + l_1 + l_2 - 1)z^{-2} + \\ + \ldots + (l_0 + l_1 + l_2 + \ldots + l_{n-1} - 1)z^{-(n-1)}$$

$l_\mu = P\{a_k = \mu\}$ for $\mu = 0, \ldots, n$ is the probability of exactly μ arrivals at any time slot; and λ is the mean arrival rate, that is, $\lambda = \sum_{\mu=1}^{n} \mu l_\mu$.

This method is computationally reliable and simple, and can be carried out by employing one of several available computational tools. The previous expression is valid for different values of the mean arrival rate(λ), even greater than 1. For the case of interest, when $\lambda < 1$, for large values of the maximum delay tolerance (T) the dropping rate will depend almost entirely on the dominant pole r_d of the above filter (largest root of $D(z)$); thus $d_r(T) \approx \alpha r_d^T$ where α is a constant as explained in [12]. In the limiting case when $\lambda = 1$, the dominant pole $r_d = 1$ and for large values of T, it follows $d_r(T) \approx \frac{\sigma_l^2}{2(T+1)}$, where σ_l^2 is the variance of the distribution of the number of arrivals.

V. The Real Fixed Frame Length TDMA (RFFL-TDMA) Scheme

Consider the RVFL-TDMA scheme in which the frame length L_k is not variable (on-demand) but fixed and equal to L_f time slots. As before, two auxiliary systems are defined:

- FL : At time t_k^g the scheduler knows about all arrivals up to time t_k^g.
- FU : At time t_k^g the scheduler knows only about the arrivals up to time $t_k^g - Re$

Let the superscripts FL, FR, and FU indicate a quantity associated with the auxiliary system FL, the RFFL-TDMA scheme and the auxiliary system FU, respectively. Similarly to Section III, it is easy to establish that: $d_r^{FL} \leq d_r^{FR} \leq d_r^{FU}$.

The auxiliary systems FL and FU are studied in order to derive tight bounds on the performance induced by the RFFL-TDMA scheme. An approach similar to the one used before to compute the dropping rate of the IVFL-TDMA scheme is employed here. The reader is referred to [12] for the complete procedure.

VI. Numerical Results and Discussion

In this section the performance of the RVFL-TDMA scheme is analyzed by employing the analytical studies presented in the previous sections. In addition to evaluating the impact of the key parameters - such as the maximum delay tolerance T, the number of applications N, the length of frame overhead, etc. - a comparison with the other two aforementioned TDMA schemes is presented.

The lower and upper bounds on the dropping rate induced under a RVFL-TDMA scheme supporting N=6 Bernoulli users are calculated by employing the auxiliary systems L and U and using the expressions derived in Section III; the per user traffic rate is equal to 0.15. The results are presented in Figure 3 as a function of the maximum delay tolerance T and for various values of (Re, In). For the (small) values of the frame overhead $(Re + In)$ considered turns out that the bounds are very tight, suggesting that

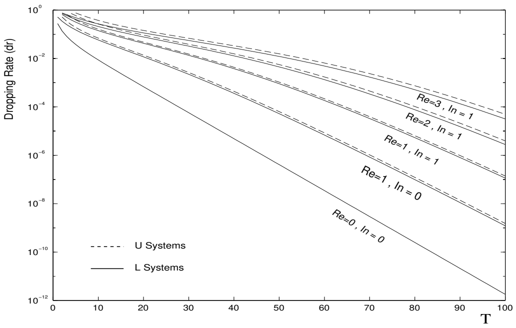

Fig. 3. Dropping rate bounds under the RVFL-TDMA scheme versus maximum delay tolerance (T) for various values of (Re, In); total traffic rate is $\lambda = 0.9$.

the results under the auxiliary system L (lower bound) can serve as a good approximation on the exact dropping rate. The results suggest that a small increase in the length of the frame overhead results in a significant increase in the dropping rate. Thus it may be worth trying to reduce the length of the frame overhead by employing mechanisms such as piggy-backing and arrival time estimation.

Figure 4 depicts the lower bound on the dropping rate under the RVFL-TDMA scheme under different traffic configurations. The results are shown as a function of T and for various values of $(Re, In) = (ov, 0)$; ov is a shorthand notation for *overhead*. Figure 4-(a) is derived for a set of $N = 6$ Bernoulli users with per user rate of 0.15 (total rate $\lambda = 0.9$); Figure 4-(b) is derived for a set of $N = 5$ Bernoulli users with per user rate of 0.20 ($\lambda = 1.0$); Figure 4-(c) is derived for a set of $N = 4$ Bernoulli users with per user rate of 0.20 ($\lambda = 0.8$); Figure 4-(d) is derived for a set of $N = 50$ Bernoulli users with per user rate of 0.016 ($\lambda = 0.8$).

From the values in Figure 4 (in log scale) the exponential decay of the dropping rate can be observed for large T. This is the case not only under zero frame overhead (as expected since this result is equal to the one obtained under the ICE-TDMA scheme of Section IV) but also in the presence of frame overhead. By considering the results in Figure 4-(c) and Figure 4-(d) for small frame overhead length it can be concluded that the induced dropping rate for $N = 50$ users (Figure 4-(d)) is significantly greater than that for $N = 4$ users (Figure 4-(c)), although $\lambda = 0.8$ in both cases. This difference may be attributed to the larger variance of the traffic for $N = 50$ and decreases as the frame overhead increases. Thus, the burstiness (variance) of the traffic in addition to the rate may impact significantly on the induced dropping rate.

The maximum number of users that can be supported under the RVFL-TDMA scheme can be determined by considering the results in Figure 5, presenting the dropping rate as a function of the number N of supported Bernoulli users and for various values of $(Re, In) = (overhead, 0)$; note that for a given (frame) overhead length, setting $Re = overhead$ and $In = 0$ will yield the lowest dropping rate for the particular TDMA scheme. The results shown in Figure 5 quantify the (negative) impact of the

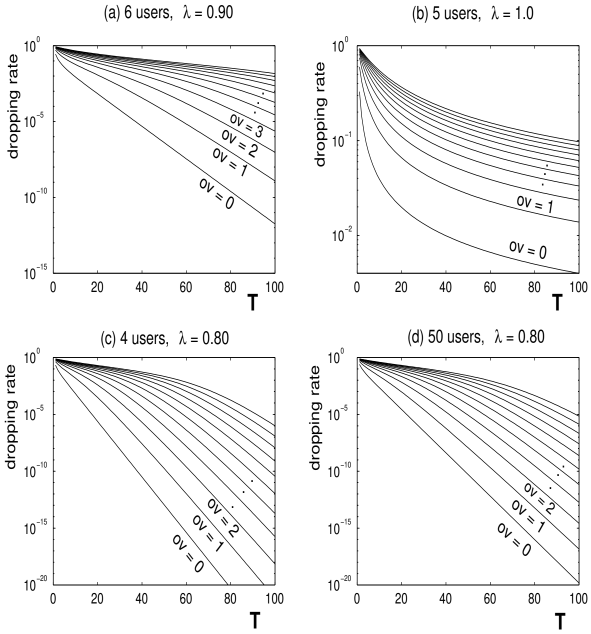

Fig. 4. Dropping rate versus maximum delay tolerance (T) for different configurations of the RVFL-TDMA scheme.

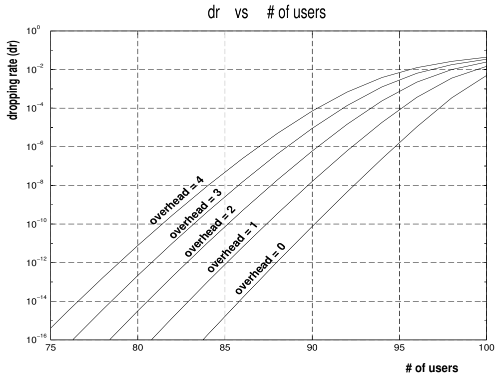

Fig. 5. Dropping rate versus number of users for the RVFL-TDMA scheme; maximum delay tolerance $T = 100$ time slots.

frame overhead on the system utilization when a certain level of QoS (dropping rate) is to be delivered; the per user rate is 0.01 and the maximum delay tolerance is 100 time slots.

If a specific cell loss probability (as opposed to system dropping rate) is desired, then the cell loss probability can be derived as the ration of the system dropping rate and total arrival rate ($\lambda = 0.01N$). For example, for a maximum cell loss probability of 10^{-12}, reducing the frame overhead length from 4 to 0 will allow to increase the number of supported users from $N \approx 78$ to $N \approx 87$, an increase of approximately equal to 11.5%.

The (system) dropping rate under the RFFL-TDMA scheme is shown in Figure 6 as a function of the (fixed) frame length L_f; the frame overhead is given by $(Re = 2, In = 0)$ and $T = 100$. Figures 6-(a) and 6-(b) are obtained for $N = 4$ ($\lambda = 0.8$) and $N = 5$ ($\lambda = 1.0$) users, respectively; the per Bernoulli user rate is 0.2. Since the bounds on the dropping rate derived from the auxiliary systems FL and FU (Section V) are very tight, only the

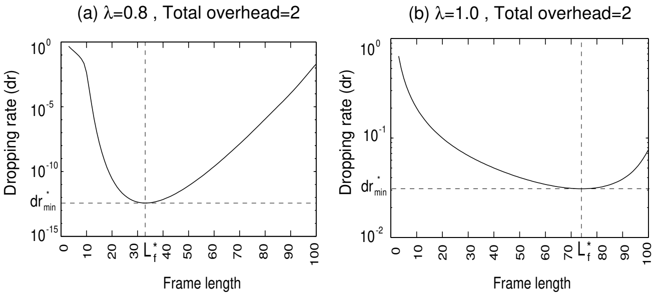

Fig. 6. Dropping rate versus frame length (fixed) for the RFFL-TDMA scheme; maximum delay tolerance $T = 100$ time slots.

lower bound is plotted. It can be observed that for a given set of supported applications there exists an optimal (fixed) frame length (namely L_f^*) minimizing the induced dropping rate. This optimal frame length is, in general, different for a different set of supported applications, as illustrated in Figure 6. It should be noted that all results under the RFFL-TDMA scheme presented below are obtained for the optimal value of the (fixed) frame length L_f^*.

Figure 7a (b) shows the dropping rate as a function of the maximum delay tolerance under both the RVFL-TDMA scheme - results also shown in Figure 4c(b) - and the RFFL-TDMA scheme. This figure illustrates the improved performance induced by a variable frame length (RVFL-TDMA) scheme compared to that of a fixed frame length (RFFL-TDMA) scheme. The values of the frame overhead considered are $(Re, In) = (1, 0)$ and $(Re, In) = (2, 0)$. The results are shown in Figure 7a and Figure 7b for $N = 4$ ($\lambda = 0.8$) and $N = 5$ ($\lambda = 1.0$) Bernoulli users, respectively. Note that the case of zero frame overhead is not considered since in this case $L_f^* = 1$ and the resulting TDMA scheme becomes equivalent to the ICE-TDMA scheme.

Figure 8 shows the dropping rate as a function of the number of users under both the RVFL-TDMA scheme (results also shown in Figure 5) and the RFFL-TDMA scheme. This figure illustrates the (positive) impact that allowing the frame length to vary on demand (instead of being fixed) has on the maximum number of admitted users, when a certain level of QoS is to be delivered. As before (Figure 5), the users are Bernoulli with a per user rate of 0.01, the maximum delay tolerance is 100 time slots, and $(Re, In) = (overhead, 0)$ (with $overhead = 1, 2, 3$). It can be noted that the gap between the schemes increases if the frame overhead length is increased or the required dropping rate is reduced (higher Quality of Service). When it is required to deliver a dropping rate of at most 10^{-16}, using the RVFL-TDMA scheme instead of the RFFL-TDMA scheme (for the same overhead) will increase the number of admitted users by up to a 10%. This gain decreases when the induced dropping rate is allowed to increase. When the required dropping rate is around 10^{-4} the difference in the number of admitted users under both schemes will be around 3%. These figures (Figures 7 and 8) quantify an important result of this paper.

The relative performance of variable frame length TDMA schemes versus fixed frame schemes is shown in Figure 9 under various traffic environments: $N = 6$ Bernoulli users each of rate 0.15 ($\lambda = 0.9$) are considered in case (a) and $N = 5$ Bernoulli users each of rate 0.20 ($\lambda = 1.0$) are considered in case (b). $N = 8$ ($N = 10$) bursty users with total rate of $\lambda = 0.8$ ($\lambda = 1.0$) are considered in case (c) (case (d)). A bursty user generates 0 or 10 cells per slot with corresponding probabilities 0.99 and 0.01, resulting in a rate of 0.1 cells/slot. Results are presented under the FF-TDMA scheme with $L_f = N$, the IVFL-TDMA (or ICE-TDMA) scheme and various RVFL$_k$-TDMA schemes, where k represents the total overhead. The IVFL-TDMA scheme is the RVFL$_k$-TDMA Scheme with $k = 0$ (no overhead).

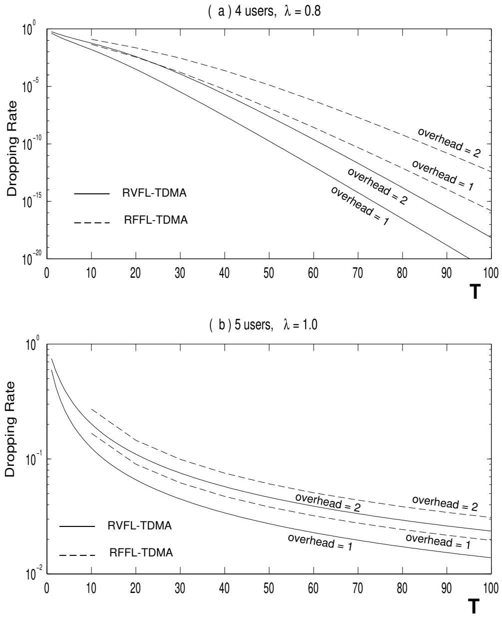

Fig. 7. Dropping rate under RVFL-TDMA and RFFL-TDMA schemes versus maximum delay tolerance (T).

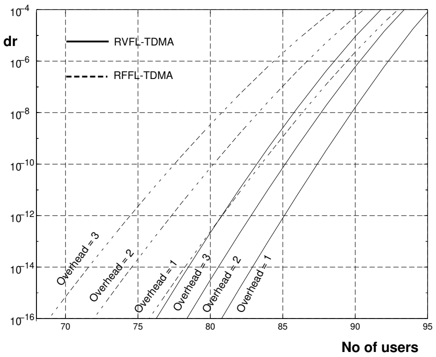

Fig. 8. Dropping rate under RVFL-TDMA and RFFL-TDMA schemes versus number of users; maximum delay tolerance $T = 100$ time slots.

The FF-TDMA scheme with $L_f = N$ models a TDMA scheme which allocates one slot to each user; no overhead is present since no communication between the users and the scheduler is necessary (similar to the TDM scheme used in the T-1 system for voice transmission). The results under this FF-TDMA scheme can be obtained by calculating the dropping rate for each user from the study of an FL system

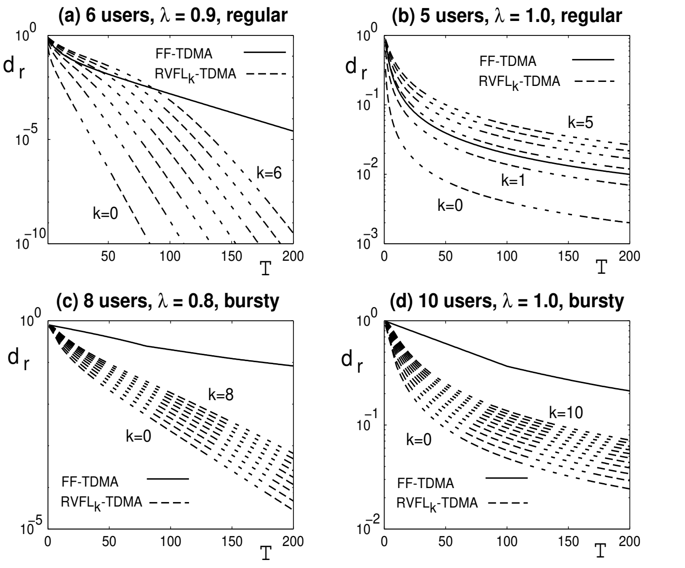

Fig. 9. Dropping rate (dr) versus maximum delay tolerance (T) under the FF-TDMA, RVFL$_k$-TDMA, and ICE-TDMA schemes for different traffic environments.

with parameters $L_f = N, Re = N - 1$ and $In = 0$ and adding the results for all users. That is, each user may be viewed as being alone in a RFFL-TDMA scheme with overhead $Re = N - 1, In = 0$.

The IVFL-TDMA (ICE-TDMA) scheme - RVFL$_k$-TDMA scheme with $k = 0$ - is the optimal scheme yielding the minimum possible system dropping rate. No slots are wasted under the IVFL-TDMA (ICE-TDMA) scheme (no pre-allocation of slots) while user/scheduler information exchange is assumed without any overhead ($k = 0$).

The RVFL$_N$-TDMA scheme (i.e. $k = N$) considered in this figure assumes a contention-free request reservation scheme which utilizes $Re = N$ full slots per frame. That is, each user has its own full slot for reservations; $In = 0$. Typically, it is expected that minislots (as opposed to slots) would be assigned to users for reservations yielding to a much smaller value of Re. The results under RVFL$_k$-TDMA schemes are obtained by employing the auxiliary system L analyzed in Section III.A.

The results shown in Figure 9 demonstrate behavior anticipated from the insight and discussions presented in this paper. No scheme outperforms the IVFL-TDMA (ICE-TDMA) scheme under any system configuration. Under bursty traffic the FF-TDMA scheme is the poorest under all values of λ and T considered. Under the (less bursty, or more regular) Bernoulli traffic and rate $\lambda = 1$ the FF-TDMA scheme outperforms some of the RVFL$_k$-TDMA schemes (for $k \geq 2$), while for lower rates ($\lambda = 0.9$) the FF-TDMA scheme is outperformed by the RVFL$_k$-TDMA scheme for small k and/or large values of T.

VII. Conclusions

In this paper the performance of the variable frame TDMA scheme has been investigated analytically and compared to other related schemes. The (minimum) dropping rate induced under the (optimal) IVFL-TDMA scheme has been derived by employing a numerical approach, while presenting the more insightful analytical derivation of this quantity under the (equivalent in this paper) ICE-TDMA scheme. The cases of the sub-optimal (non-zero overhead) RVFL-TDMA and (fixed frame) RFFL-TDMA schemes has been considered as well and their performance has been compared to that of the IVFL-TDMA scheme.

A number of interesting conclusions have been drawn from the insightful consideration of these schemes and the presented numerical results. It has been observed that for the arrival rates of interest the frame overhead impacts significantly on the achievable dropping rate, thus, every effort should be made to reduce the overhead. A comparison between the RVFL-TDMA and the RFFL-TDMA schemes shows that for the same amount of overhead, the RFFL-TDMA scheme induces significantly larger system dropping rate. A comparison between the RVFL-TDMA with contention free reservation period and the FF-TDMA schemes shows that no one scheme outperforms the other in all cases; the formulas developed in this paper can be used to identify the best one for a particular case. Both schemes are far from the optimal, so there is room for significant improvement using techniques as for example piggy-backing of the arrival information, or arrival time estimation.

References

[1] I.M. Leslie et al., "*ATM Everywhere'*" , IEEE Networks, pp 40-46, March 1993.
[2] B. Walke et al. , "*Wireless ATM: Air Interface and Network Protocols of the Mobile Broadband System*", IEEE Personal Communications Mag., 3(4), pp. 50-56, August 1996.
[3] D. Raychaudhuri and N. Wilson. "*ATM-Based Transport Architecture for Multiservices Wireless Personal Communication Network*". IEEE Journal of Selected Areas Communications, 12(8):1401-1414, October 1994.
[4] K.Y. Eng, M. Karol, and M. Veeraraghavan. "*A Wireless Broadband Ad-Hoc ATM Local Area Network*". ACM Wireless Networks Journal, 1(2), December 1995.
[5] P. Mermelstien, A. Jalali, and H. Leib. "*Integrating Services on Wireless Multiple Access Networks*". In Proceedings of ICC, pages 863-867, March 1993.
[6] D. Petras et al. "*MAC Protocol for Wireless ATM: Contention Free versus Contention Based Transmission reservation Request*". In proceedings of PIMRC, Taipei, Taiwan, October 1996.
[7] M. Karol, Z. Liu, and K. Eng. "*An Efficient Demand-Assignment Multiple Access Protocol for Wireless Packet (ATM) Networks*". Wireless Networks, 1(4):267-279, December 1995.
[8] A. Mahmoud, D. Falconer, and S. Mahmoud. "*A Multiple Access Scheme for Wireless Access to a Broadband ATM LAN Based on Polling and Sectored Antennas*". IEEE Journal of Selected Areas Communications, 14(4), May 1996.
[9] C. Chang, K. Chen, and M. You. "*Guaranteed Quality of Service Wireless Access to ATM networks*". In ICC, Dallas, Tx. June 1996.
[10] N. Passas, L. Merakos, and D. Skyrianoglou. "*MAC Protocol and Traffic Scheduling for Wireless ATM Networks*". To appear in ACM Mobile Networks and Applications Journal, special issue on Wireless LANs, 1998.
[11] S.S. Panwar et. al., "*Optimal Scheduling Policies for a Class of Queues with Customer Deadlines to the Beginning of Service*" , Journal of the ACM, pp. 832-844, 1988.
[12] C. Santiváñez, I. Stavrakakis."*Evaluation and Comparative Study of TDMA Schemes for Real Time Application Support*". CDSP Report TR-CDSP-98-48, CDSP Center, Department of Electrical and Computer Engineering, Northeastern University, Boston, MA. Available in http://www.cdsp.neu.edu/info/students/cesar/cdsp_report.ps .