Analysis of Non-Gated Integrated Services TDM

Ioannis Stavrakakis and Sophia Tsakiridou

Abstract— **A very flexible analysis approach is developed for the exact analysis of a non-gated TDM scheme, supporting Time Critical (*TC*) and Time Non Critical (*TNC*) packetized information. Correlations in the *TNC* traffic and spreading of the *TC* traffic over the entire frame can be accommodated, unlike in the past work.**

I. INTRODUCTION

A Time Division Multiplexing (TDM) protocol is considered in this paper for the accommodation of Time Critical (*TC*) and Time Non Critical (*TNC*) sources (traffic) in an Integrated Services Digital Network (ISDN). Details regarding the motivation of this work, related past studies and the developed analysis approach may be found in [1]. *TC* traffic may be associated with real time applications such as voice and video. *TNC* traffic may represent computer data traffic.

The TDM frame (of length T) is divided into two parts: the *TC* subframe (of length T_{TC}) and the *TNC* subframe (of length T_{TNC}), which are used for the transmission of *TC* and *TNC* information, respectively (Fig. 1). Under the fixed boundary TDM policy, the two subframes assume constant values throughout the operation of the system. Under the movable boundary TDM policy, *TNC* sources are allowed to transmit over the unused portion of the *TC* subframe.

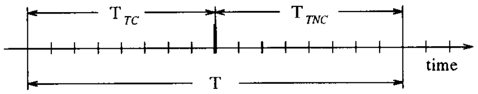

Fig. 1. The *TC* and the *TNC* subframes of the TDM frame.

The movable boundary TDM protocol has been adopted for the voice/data integration and has been analyzed in the past both in continuous and discrete time, [1]. The most relevant studies in discrete-time have been based on the generating function approach, [2], [3], [4], [5], [6], [7]. The discrete-time, movable boundary TDM system considered in this paper is different from those studied in the past in at least two aspects. First, the *TNC* traffic is assumed to be correlated. Second, the *TC* traffic is not necessarily accommodated in a subframe of contiguous slots, but it may be spread over the entire frame according to any pattern. The non-gated service policy for the TNC sources is assumed, unlike all previous work except from [6], [7]. Furthermore, the developed flexible analysis approach is capable of handling, in a unified way, TDM schemes under a variety of conditions.

II. SYSTEM MODEL AND ANALYSIS APPROACH

Information is organized in packets of fixed size with transmission time equal to one slot. The M TC sources are assumed to be identical, each delivering one packet every T slots, when active (state 1). No packet is delivered when the *TC* source is inactive (state 0). A two state Markov model is adopted for the *TC* source (Fig. 2). *TC* packets must be transmitted in the first frame following their generation time instant, otherwise they are useless and they are dropped.

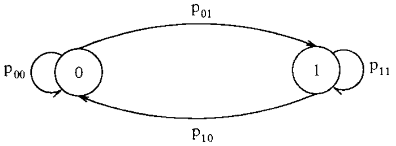

Fig. 2. The Markov model for the *TC* sources; the burstiness coefficient γ of this two-state Markov source is defined by $\gamma = p_{11} - p_{01}$.

Unlike all related past work, the *TNC* traffic is assumed to be correlated; it is modeled as a Markov Modulated Generalized Bernoulli (MMGB) process, [1]. According to this process, packet arrivals are governed by an underlying Markov Chain (with some state space S'). Transitions between states of the chain occur at slot boundaries. The packet arrival process is determined in terms of a probabilistic mapping $a'(\cdot) : S' \rightarrow \{0, 1, \ldots, R'\}$, where R', $R' < \infty$, is the maximum number of packets delivered per slot by the MMGB process.

To simplify the notation and discussion in this section, the movable boundary policy is adopted, the relation $T_{TC} \leq T$ is assumed to hold (no blocking of TC traffic), and TC packets are assumed to be transmitted over the first R_i (contiguous) slots of the i-th frame, at the beginning of which R_i TC sources are active. In Section III it is outlined how the developed flexible analysis approach can accommodate TDM systems under other conditions.

The TDM system can be described in terms of a discrete-time queueing system with interruptions, as shown in Fig. 3. The buffer capacity of the *TNC* (*TC*) queue is assumed to be infinite (equal to $2T_{TC}$). To reflect the TDM

Paper approved by J.W. Wong, the Editor for Wide Area Networks of the IEEE Communications Society. Manuscript received March 2, 1992; revised November 2, 1993 and October 12, 1993. This work was supported in part by the National Science Foundation under Grant NCR-9011962. This paper was presented in part at the IEEE INFOCOM'92 Conference, May 4-8, 1992, Florence, Italy.

I. Stavrakakis is currently with the Department of Electrical and Computer Engineering of Northeastern University, Boston, Massachusetts 02115, USA, on leave from the University of Vermont.

S. Tsakiridou is with the Department of Computer Science and Electrical Engineering of the University of Vermont, Burlington, Vermont 05405-0156, USA.

IEEE Log Number 9410862.

0090-6778/95$4.00 © 1995 IEEE

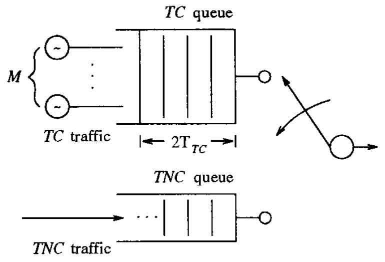

Fig. 3. The discrete-time queueing model for the adopted TDM policy.

policy, the server is assumed to switch to the TC queue at the beginning of a frame and serve all TC packets found upon switching (at most T_{TC}); then, it switches to the TNC queue and serves the TNC traffic according to the First-In First-Out (FIFO) service discipline. At the end of each frame, the TC packets which have been queued for at least T slots are dropped. Notice that a TNC packet is transmitted as soon as it reaches the head of the TNC queue provided that the server is available (non-gated discipline).

The performance analysis of the TC queue is straightforward, [1]. The analytically challenging problem associated with the TDM system is that of the evaluation of the queueing behavior of the TNC packets, which is affected by the activity in the TC queue. Fortunately, the TDM policy is such that it presents a one direction interference on the TNC queue coming from the TC queue; the behavior of the TC queue is, on the other hand, independent from the activity at the TNC queue. This observation allows for modeling the interference from the TC queue on the TNC queue as an independent—from anything associated with the TNC queue—process $\{\bar{R}_j\}_{j\geq 0}$, as it is explained below. The TNC queue is studied then, under the TNC packet traffic and the interfering process $\{\bar{R}_j\}_{j\geq 0}$ which captures the TDM policy. This approach is new and different from previous ones developed for the study of integrated TDM systems. Non-gated policies considered in the past, some new policies, as well as correlations in the TNC traffic (considered here for the first time), can be analyzed in a unified way, by selecting the appropriate interfering process $\{\bar{R}_j\}_{j\geq 0}$.

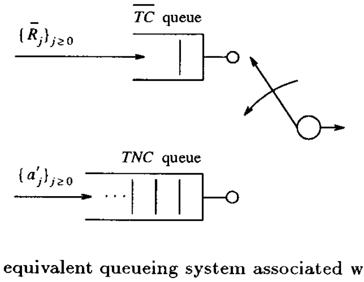

Fig. 4. The equivalent queueing system associated with the TNC queue.

The queueing behavior of the TNC packets of the TDM system (Fig. 3) is identical to that of the TNC packets in the system shown in Fig. 4. Whenever the server in Fig. 3 is not available to TNC packets (first $R_{\hat{\imath}}$ slots of the $\hat{\imath}$-th frame), the server in Fig. 4 serves packets in the $\overline{TC}$ queue. The previous can be implemented by constructing a proper arrival process to $\overline{TC}$ ($\{\bar{R}_j\}_{j\geq 0}$) and adopting a proper service policy as follows. The interference process $\{\bar{R}_j\}_{j\geq 0}$ is defined by: $\bar{R}_j = 1$ if $0 \leq j \bmod \mathrm{T} \leq R_{\hat{\imath}} - 1$, $\hat{\imath} = \lfloor j/\mathrm{T} \rfloor$, and $\bar{R}_j = 0$ otherwise; $\lfloor x \rfloor$ denotes the integer part of x. Notice that one packet is delivered by $\{\bar{R}_j\}_{j\geq 0}$ at the j-th slot, if this slot is in the first $R_{\hat{\imath}}$ positions of the $\hat{\imath}$-th frame. A realization of $\{\bar{R}_j\}_{j\geq 0}$ is shown in Fig. 5. The server switches to the TNC queue only if the $\overline{TC}$ queue is empty; it switches back to the $\overline{TC}$ queue, as soon as it becomes non-empty.

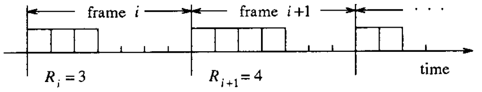

Fig. 5. A realization of $\{\bar{R}_j\}_{j\geq 0}$ for $i\mathrm{T} \leq j \leq (i+2)\mathrm{T}+3$ and $\mathrm{T} = 7$.

The behavior of the TNC queue (Fig. 4) can be studied by considering the equivalent system without service interruptions (Fig. 6) and assuming HoL priority policy for the packets delivered by $\{\bar{R}_j\}_{j\geq 0}$. Let D_{TNC} ($D_{\overline{TC}}$) denote the mean packet delay of the TNC ($\overline{TC}$) queue; $D_{\overline{TC}} = 1$ due to the adopted service policy and the fact that $\{\bar{R}_j\}_{j\geq 0}$ may deliver at most one packet per slot. If D_{FIFO} denotes the mean packet delay induced under the FIFO policy for the queueing system in Fig. 6, then D_{TNC} can be obtained from (conservation law): be obtained from (conservation law): $D_{FIFO} = \dfrac{\lambda_{TNC} D_{TNC} + \lambda_{\overline{TC}} D_{\overline{TC}}}{\lambda_{TNC} + \lambda_{\overline{TC}}}$, where λ_{TNC} ($\lambda_{\overline{TC}}$) is the input traffic rate to the TNC ($\overline{TC}$) queue.

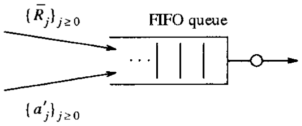

Fig. 6. The equivalent queueing system for the study of the TNC queue.

In [1] it is shown that $\{\bar{R}_j\}_{j\geq 0}$ can be described in terms of an appropriate MMGB model based on a two-dimensional underlying Markov Chain $\{R_{\hat{\imath}}, L_j\}_{j\geq 0}$ defined at slot boundaries; $\hat{\imath} = \lfloor j/\mathrm{T} \rfloor$ and $L_j = j \bmod \mathrm{T} + 1$. $R_{\hat{\imath}}$ is a Markov Chain embedded at frame boundaries, easily constructed in terms of the M independent Markov Chains for the TC sources, [1]. The probabilistic mapping is given by: $\bar{R}_j = \bar{a}\{(R_{\hat{\imath}}, L_j)\} = 1$ with probability 1, if $L_j \leq R_{\hat{\imath}}$ and $\bar{R}_j = \bar{a}\{(R_{\hat{\imath}}, L_j)\} = 0$ with probability 1, otherwise; it is defined on $S_R \times S_L = \{0, 1, \ldots, M\} \times \{1, 2, \ldots, \mathrm{T}\}$. If $R_{\hat{\imath}}$ TC users are active in the $\hat{\imath}$-th frame, then the above mapping will generate one packet over the first $R_{\hat{\imath}}$ slots of this

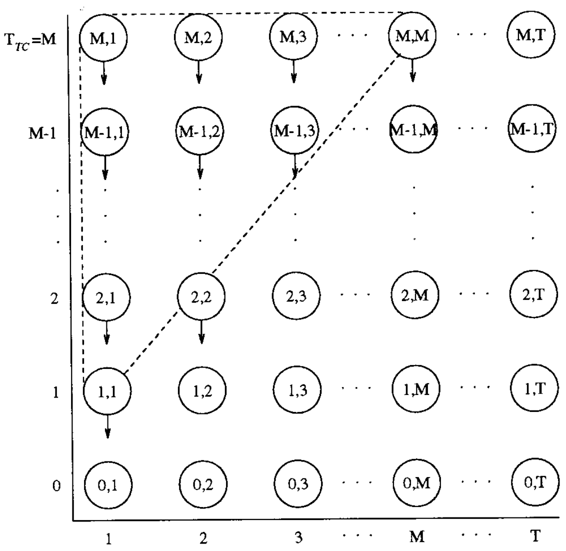

Fig. 7. The state space $S_R \times S_L$; the arrows indicate a packet generating state ($M = T_{TC} \leq T$).

frame. This is illustrated with the arrows originating from the appropriate states in Fig. 7. The packet generating states are contained in the triangle defined by the states $(1,1)$, $(T_{TC},1)$ and (T_{TC},T_{TC}). The FIFO system under MMGB arrivals has been analyzed in [8], [9]; D_{FIFO} can be calculated, [1], and D_{TNC} is obtained by invoking the conservation law.

III. NUMERICAL RESULTS AND COMMENTS

The non-gated (for *TNC* traffic) integrated services TDM considered in Section II has the following characteristics: it is a movable boundary scheme [a]; it is non-blocking for the *TC* traffic ($M = T_{TC} \leq T$) [b]; it assumes correlated *TC* sources [c] and correlated *TNC* traffic [d]; *TC* sources generate one packet when active [e]; *TC* packets are transmitted over contiguous slots [f]. In [1] it is shown how characteristics [a]-[f] can be modified and the analysis approach be still applicable. The new characteristics are captured by modifying the state space of the underlying Markov Chain (Fig. 7) and/or the associated probabilities, and/or by moving around the arrows in Fig. 7 to capture the proper probabilistic mapping; the existence of an arrow with probability less than one is assumed under modification of [e], to allow for non-certain packet generation from an active state. The spreading of *TC* packets over a frame according to some pattern is captured through the probabilistic mapping of $\{\bar{R}_j\}_{j\geq 0}$ by moving the arrows in Fig. 7 to the appropriate packet generating states determined by the spreading pattern. An algorithm for near-uniform spreading may be found in [1].

Curve (i) in Fig. 8 presents D_{TNC} under correlated (packetized voice) *TC* sources and correlated *TNC* traffic; correlation is measured in terms of the burstiness coefficient γ, defined by $\gamma = p_{11} - p_{01}$, where p_{ij} are the transition probabilities from state i to state j of the 2-state Markov model (Fig. 2). The source mean sojourn times in states 0 and 1 are increasing functions of γ. Thus, γ determines the 'burst' length of the traffic generated by the source for a given packet generation rate; the larger the value of γ, the higher the burstiness. No past work is applicable under correlated *TNC* traffic. Curve (ii) is derived under the i.i.d. model for the *TNC* traffic ($\gamma_{TNC} = 0$). This case has also been considered in [6], [7]; due to the enormous numerical complexity, bounds are derived and they are computed only for a small range of (M,T). Curves (i) and (ii) indicate that D_{TNC} increases with the burstiness of the *TNC* traffic. Curves (iii) and (iv) correspond to (i) and (ii), respectively, assuming uncorrelated *TC* sources (γ_{TC}=0).

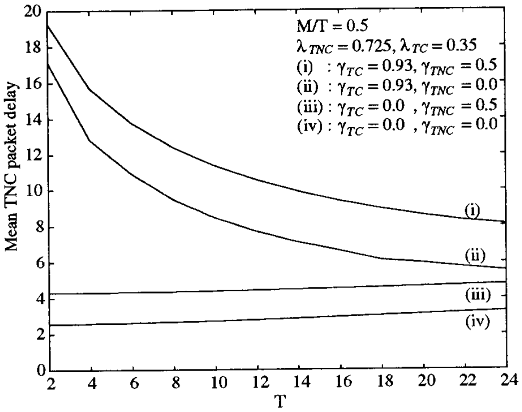

Fig. 8. Mean *TNC* packet delay results. A two-state MMGB model (Fig. 2) generating one packet from state 1 with probability one is adopted for the *TNC* traffic. γ_{TC} (γ_{TNC}) denotes the burstiness of a *TC* source (the *TNC* traffic); $T_{TC} = M$.

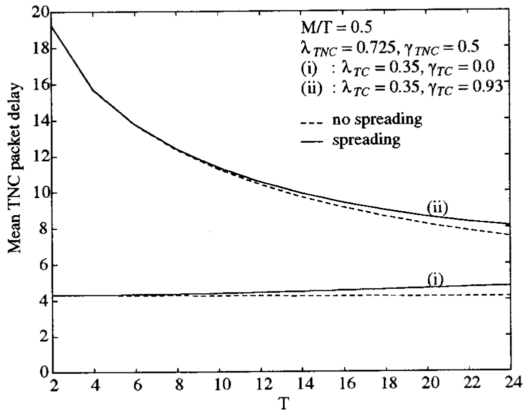

Fig. 9. Mean *TNC* packet delay results under near-uniform spreading of the *TC* traffic, [1].

The case of the uniform spreading of the *TC* packets over the frame has not been considered in the past. Fig. 9 presents results under the (near-uniform) spreading policy in [1]. Notice that the positive impact of the spreading policy on the induced D_{TNC} becomes more pronounced as the frame size T increases, since 'more' spreading is possible. Finally, the TNC queue occupancy distribution has also been derived in [1] by applying matrix-analytic techniques.

The numerical complexity of the developed approach (order of T^2M^2) seems to be similar to that in [5], where the gated policy for the TNC traffic is assumed. Simulation results for large values of T (where the numerical complexity is large) indicate that all curves practically converge. Thus, i.i.d. approximations—which have been observed to require less computation time—could be adopted when T is large.

REFERENCES

[1] I. Stavrakakis and S. Tsakiridou, "Analysis of non-gated integrated services TDM," Research Report CSEE/93/02-01, CSEE Dept., University of Vermont, Feb. 1993.

[2] M.J. Fischer and T.C. Harris, "A model for evaluating the performance of an integrated circuit-and packet-switched multiplex structure," *IEEE Trans. Commun.*, vol. COM-24, pp. 198–202, 1976.

[3] K. Sriram, K.K. Varshney, and J.G. Shanthikumar, "Discrete-time analysis of integrated voice/data multiplexers with and without speech activity detectors," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 1124–1132, 1983.

[4] San-Qi Li and J.W. Mark, "Performance of voice data integration on a TDM system," *IEEE Trans. Commun.*, vol. COM-33, pp. 1265–1273, 1985.

[5] O.W. Yang and J.W. Mark, "Queueing analysis of an integrated services TDM system using a matrix-analytic method," *IEEE J. Select. Areas Commun.*, vol. SAC-9, pp. 88–94, 1991.

[6] R.F. Chang and V.O.K. Li, "Analysis of packet delay for an integrated voice-data system," in *Proc. IEEE GLOBECOM'88*, Hollywood, FL, Nov. 1988, pp. 1095–1100.

[7] Z. Zhang and I. Rubin, "Bounds on the mean system-size and delay for a movable-boundary integrated circuit and packet switched communications channel," in *Proc. IEEE INFOCOM'90*, San Francisco, CA, June 1990, pp. 866–873.

[8] I. Stavrakakis, "Analysis of a statistical multiplexer under a general input traffic model," in *Proc. IEEE INFOCOM'90*, San Francisco, CA, June 1990, pp. 1220–1225.

[9] I. Stavrakakis, "Efficient modeling of merging and splitting processes in large networking structures," *IEEE J. Select. Areas Commun.*, vol. SAC-9, pp. 1336–1347, 1991.