

LOSS VOLUME IN CONTINUOUS FLOW MODELS: FAST SIMULATION AND SENSITIVITY ANALYSIS VIA IPA

YORAI WARDI[†], BENJAMIN MELAMED[‡]

[†]Georgia Institute of Technology, School of Electrical and Computer Engineering, Atlanta, GA 30332, Yorai.Wardi@ece.gatech.edu

[‡]Rutgers University, Faculty of Management, Department of Management Science and Information Systems, Piscataway, NJ 08854, melamed@rbs.rutgers.edu

Abstract. This paper defines a class of fluid-flow models, called Continuous Flow Models (CFMs), representable as DEDS (Discrete Event Dynamic System) models. The CFM class is motivated by emerging high-speed packet-based telecommunications networks for which traditional queueing simulations at the packet level are prohibitively costly in time and space, or simply infeasible. In contrast, CFM-based fluid-flow networks hold the promise of fast simulation for applications ranging from network design to network control. The paper studies the loss volume metric of a basic CFM in some detail. This metric is easily converted to loss probabilities – an important ingredient in quality of service (QoS) metrics for modern telecommunications networks. The paper further performs sensitivity analysis of CFMs via IPA derivatives of the loss volume as function of buffer size, as well as service rate and arrival rate parameters. Simple formulas for these derivatives are derived and shown to be amenable to real-time computation. The formulas have a broad applicability due to their nonparametric (distribution-free) nature, a fact that makes them potentially suitable for real-time control applications in telecommunications networks.

Key words. CFM, Continuous Flow Models, High-Speed Networks, IPA, Loss Metrics.

Topics addressed. Discrete event systems.

1 Introduction

Loss probabilities are important performance measures in real-time applications of packet-switched telecommunications networks. Quality of service (QoS) requirements are often stated in terms of maximal allowed loss probabilities, e.g., variable bit-rate (VBR) service in ATM networks. Packet loss can be ameliorated by allocating sufficient network resources, either prior to operation as in the case of design, or in real time, as in the case of control. In either case, it may be necessary to estimate the loss probability or loss volume, and their derivatives with respect to various network parameters.

Barring special situations or simplifying modeling assumptions, simulation tools may have to be employed for estimating loss-related performance measures. Unfortunately, discrete-event simulators, based on traditional queueing network models, may require prohibitive amounts of computing time. The reason is that *events* are typically associated with arrivals, processing or departures of packets at network nodes. But in current and emerging packet networks (e.g., ATM networks), operating at gigabit-per-second rates to transport small packets, such network simulations can easily call for millions of events per second (per node), thus rendering large-network simulation impractical. To get around this problem, an alternative modeling scheme has been considered, based on fluid-flow rather than on the discrete character of individual packets. When traffic flow rates are piecewise-constant functions of time, it is easy to represent a fluid-flow model as a DEDS (Discrete Event Dynamic System), where events largely model rate changes. Since these events typically occur far less frequently than packet-related events in traditional queueing models, fluid-based simulation are far more practical. Additionally, the CFM class is amenable to IPA-based sensitivity analysis, unlike traditional queueing systems.

First employed in [1] and subsequently in [3] in the analysis of multiplexed data streams, fluid-flow models were later used for simulation and sensitivity analysis of loss-related measures [5, 4]. A general paradigm for fast simulation of fluid-flow systems, called *Continuous Flow Models* (CFM), was devised in [6], which further established the unbiasedness of associated *Infinitesimal Perturbation Analysis* (IPA) derivative esti-

matoms [2] in a general network context.

This paper is concerned with the loss volume in a single CFM node, and its sensitivity analysis with respect to various CFM parameters. It shows how to implement fast CFM simulations via a Lindley-like equation, and it develops some IPA estimators for sensitivity analysis of CFM systems. These estimators can be computed by fairly simple recursive equations, which in some cases are *nonparametric* (that is, they require no assumptions on the distributions of the underlying processes.) Consequently, such estimates can be computed in real time from either simulation runs or observations of a real-life system (network, node), in which case they can be used in control environments. The formulas can be used in various telecommunications applications, such as network design, admission and congestion control, service rate allocation, and buffer size allocation.

The rest of this paper is organized as follows. Section 2 defines loss measures in a basic single-server CFM, for which it derives a simple Lindley-like equation for fluid volume in the buffer. Section 3 addresses derivative estimation, and derives fast formulas for its computation. Finally Section 4 concludes the paper.

2 Basic CFMs and Fluid Loss

A single-node CFM consists of a buffer that holds fluid, and a server that discharges fluid from the buffer. At time t , fluid flows in at rate $\alpha(t)$, and is discharged by the server at rate $\beta(t)$. The buffer workload (contents) is denoted by $x(t)$, and the buffer capacity by $c(t)$. The fluid outflow rate from the server, and the loss (spillover) rate due to finite buffer space are denoted by $\delta(t)$ and $\gamma(t)$, respectively. Note that the buffer capacity, $c(t)$, can be time dependent; this modeling wrinkle is inspired by ATM networks, where the buffer allocated to a service class may depend on demands from service classes of higher priority.

Let the time variable, t , be confined to a prescribed interval $[0, T]$, and suppose that $\{\alpha(t)\}$, $\{\beta(t)\}$, etc. are random processes, over some probability space (Ω, \mathcal{F}, P) . In the sequel, we will be concerned with sample-path properties of the above processes, so all statements will be meant to hold "with probability one". It is further assumed that initially the buffer is empty,

i.e., $x(0) = 0$, and that the sample paths $\alpha(\cdot)$, $\beta(\cdot)$ and $\dot{c}(\cdot)$ are piecewise continuously differentiable, each having a finite number of discontinuity points in $[0, T]$. Note that a dot above a function symbol signifies a derivative with respect to time t , while a dot replacing a time argument signifies a sample path.

Following [6], the processes $\{\alpha(t)\}$, $\{\beta(t)\}$ and $\{\dot{c}(t)\}$ are referred to as the *defining processes*, while the processes $\{x(t)\}$, $\{\delta(t)\}$ and $\{\gamma(t)\}$ are referred to as the *derived processes*. The reason is that the latter three processes can be computed from the former three processes. To see this observe [6] that

$$\frac{dx(t)}{dt^+} = \begin{cases} 0, & \text{if } x(t) = 0 \text{ and } \alpha(t) \leq \beta(t) \\ \dot{c}(t), & \text{if } x(t) = c(t) \text{ and } \alpha(t) - \beta(t) \geq \dot{c}(t) \\ \alpha(t) - \beta(t), & \text{otherwise;} \end{cases} \quad (1)$$

$$\delta(t) = \begin{cases} \beta(t), & \text{if } x(t) > 0 \\ \alpha(t), & \text{if } x(t) = 0; \end{cases} \quad (2)$$

$$\gamma(t) = \begin{cases} \alpha(t) - \beta(t) - \dot{c}(t), & \text{if } x(t) = c(t) \text{ and } \alpha(t) - \beta(t) \geq \dot{c}(t) \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Observe that Eq. (1) constitutes a dynamical system, whose state is the workload $x(t)$. Furthermore, (1) has a unique solution in view of the assumption $x(0) = 0$.

The performance measure of interest in this paper is the loss volume in the interval $[0, T]$, defined by

$$L(T) = \int_0^T \gamma(\tau) d\tau. \quad (4)$$

This time-average measure can be normalized to yield the loss probability

$$P_{L(T)} = \frac{L(T)}{\int_0^T \alpha(\tau) d\tau} \quad (5)$$

in the interval $[0, T]$, but to simplify the development we will treat the loss volume instead.

The CFM of Eqs. (1) – (3) can serve as a building block for constructing general CFM networks and their associated simulation programs. A CFM network is constructed as an interconnected system of such CFMs, with specific rules for routing the flow processes among them. For this reason, the single-server CFM model is called the *basic*

CFM. For examples and further discussion, including the modeling of multi-class flow systems using basic CFMs, see [6].

The important case of piecewise-constant flow rates facilitates the application of the discrete-event simulation paradigm to model CFM-based systems. To see that, consider the basic CFM, and suppose that realizations of the defining processes, $\{\alpha(t)\}$, $\{\beta(t)\}$ and $\{\dot{c}(t)\}$, are piecewise constant, with finite numbers of jump points in the interval $[0, T]$. For simplicity, assume that the buffer capacity is always positive. From Eqs. (1) – (3) it follows that the process $\{c(t)\}$ and the derived process $\{x(t)\}$ are piecewise linear and continuous, while the derived processes $\{\delta(t)\}$ and $\{\gamma(t)\}$ are piecewise constant. We define an *event* to be any jump in either one of the defining processes $\{\alpha(t)\}$, $\{\beta(t)\}$, or $\{\dot{c}(t)\}$.

The computation of the state $x(t)$ at event times is aided by the fact that the state equation (1) has a structure similar to a Lindley equation for a queue. The derivation will follow [6]. Suppose there are N events in the interval $[0, T]$, whose occurrence times are $\{t_i\}_{i=1}^N$ in increasing order, and define $t_0 = 0$ and $t_{N+1} = T$. Let α_i , β_i , and \dot{c}_i denote the respective values of $\alpha(t)$, $\beta(t)$, and $\dot{c}(t)$ in the interval $[t_i, t_{i+1})$ (recall that these processes have constant values there, by assumption). Similarly, denote $x_i = x(t_i)$ and $c_i = c(t_i)$. Finally, let

$$L_i = \int_{t_i}^{t_{i+1}} \gamma(\tau) d\tau \quad (6)$$

be the partial loss volume in the interval $[t_i, t_{i+1})$. Clearly, by Eq. (4), one has

$$L(T) = \sum_{i=0}^N L_i. \quad (7)$$

As to the computation of the state, it was shown in [6] that,

$$x_{i+1} = \min\{\max\{x_i + [\alpha_i - \beta_i][t_{i+1} - t_i], 0\}, c_{i+1}\}. \quad (8)$$

The computation of L_i assumes a simple form, shown below, which together with the recursive equation (8) provides an especially efficient way for computing the loss volume $L(T)$. The proposition below makes use of the difference operator $\Delta t_i = t_{i+1} - t_i$.

Proposition 2.1 The partial loss volume, L_i ,

has the representation

$$L_i = \begin{cases} [\alpha_i - \beta_i]\Delta t_i + x_i - c_{i+1}, & \text{if } x_{i+1} = c_{i+1} \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Proof. Since the processes $\{x(t)\}$ and $\{c(t)\}$ are linear in the interval $[t_i, t_{i+1}]$, the buffer can become full in part of such an interval if and only if $x_{i+1} = c_{i+1}$. On the other hand, from Eq. (3), $L_i = 0$ if $x_{i+1} < c_{i+1}$.

Next, suppose that $x_{i+1} = c_{i+1}$, and let time $t^* = \min\{\tau \geq t_i : x(\tau) = c(\tau)\}$ be the first time point in the interval $[t_i, t_{i+1}]$ at which the buffer becomes full. By Eq. (3),

$$L_i = [\alpha_i - \beta_i - \dot{c}_i][t_{i+1} - t^*]. \quad (10)$$

Now, since the buffer is neither full nor empty in the interval (t_i, t^*) , Eq. (1) implies that

$$x(t^*) = x_i + [\alpha_i - \beta_i][t^* - t_i]. \quad (11)$$

Adding Eqs. (10) and (11), we obtain

$$L_i + x(t^*) = [\alpha_i - \beta_i]\Delta t_i + x_i - \dot{c}_i[t_{i+1} - t^*]. \quad (12)$$

But $x(t^*) = c(t^*)$, because the buffer becomes full at time t^* , while $\dot{c}_i[t_{i+1} - t^*] = c_{i+1} - c(t^*)$, whence (12) becomes

$$L_i = [\alpha_i - \beta_i]\Delta t_i + x_i - c_{i+1},$$

from which Eq. (9) follows. \square

3 Sensitivity Analysis

Let $\theta \in \Theta \subset \mathbb{R}$ be a parameter of the defining processes $\{\alpha(\theta; t)\}$, $\{\beta(\theta; t)\}$ and $\{\dot{c}(\theta; t)\}$, whose dependence on the parameter θ is indicated by the notation. Consequently, the derived processes, $\{x(\theta; t)\}$, $\{\delta(\theta; t)\}$ and $\{\gamma(\theta; t)\}$, also depend on θ . The loss volume is viewed as a function of θ , and has the form

$$L(\theta; T) = \int_0^T \gamma(\theta; \tau) d\tau. \quad (13)$$

This section will be mainly concerned with the derivative $L'(\theta; T)$, henceforth assumed to exist with probability 1. Throughout the paper, the prime operator will denote a derivative with respect to θ ; in contrast, recall that the dot operator over a function symbol denotes a derivative with respect to time t .

Suppose that for every $\theta \in \Theta$ the processes $\{\alpha(\theta; t)\}$, $\{\beta(\theta; t)\}$ and $\{\dot{c}(\theta; t)\}$ are piecewise constant, and have each a finite number of jump points in $[0, T]$; these jumps correspond to system events. Suppose further that the dependence of these processes on θ is via their jump times or their constant values between consecutive jump times. Accordingly, for $i = 1, \dots, N(\theta)$, let $t_i(\theta)$ be the event times in increasing order, with $t_0(\theta) = 0$ and $t_{N(\theta)+1}(\theta) = T$, and let $\alpha_i(\theta)$, $\beta_i(\theta)$ and $\dot{c}_i(\theta)$ denote the constant values of $\{\alpha(\theta; t)\}$, $\{\beta(\theta; t)\}$ and $\{\dot{c}(\theta; t)\}$, respectively, in the interval $(t_i(\theta), t_{i+1}(\theta))$. We shall also use the shorthand notation $\Delta t_i(\theta) = t_{i+1}(\theta) - t_i(\theta)$, $x_i(\theta) = x(\theta; t_i(\theta))$, and $c_i(\theta) = c(\theta; t_i(\theta))$ as an extension of the notation in Section 2 to quantities that depend on θ .

We next derive an expression for the derivative $L'(\theta; T)$ under a general setting, and then specialize it to a number of cases. To guarantee the existence of the derivatives in the sequel, we make the following assumption.

Assumption 3.1 (i) For every $\theta \in \Theta$, the sample paths $\alpha_i(\theta; \cdot)$, $\beta_i(\theta; \cdot)$, $\dot{c}_i(\theta; \cdot)$ and $t_i(\theta; \cdot)$, $i = 1, \dots, N(\theta)$, are continuously differentiable in some open interval containing θ , with probability 1. (ii) With probability 1, no event occurs at a time the buffer becomes full or empty. \square

Assumption 3.1 makes it straightforward to show the existence of the loss-related derivatives in the sequel, in light of Eq. (9) and the forthcoming development. For further discussion, see [6].

To start with the general case, let θ be a parameter of either $\alpha_i(\theta)$, $\beta_i(\theta)$, $\dot{c}_i(\theta)$ or $t_i(\theta)$, for $i = 1, \dots, N(\theta)$, and define the index sets

$$\begin{aligned} E &= \{0 \leq i \leq N(\theta)+1 : x_i(\theta) = 0\} \\ F &= \{0 \leq i \leq N(\theta)+1 : x_i(\theta) = c_i(\theta)\}. \end{aligned}$$

The interpretation of these sets is straightforward: $i \in E$ if and only if the buffer becomes empty during the interval $(t_{i-1}(\theta), t_i(\theta)]$, while $i \in F$ if and only if the buffer becomes full during that interval. Finally, for $i = 1, \dots, N$, define

$$k(i) = \max\{k \leq i : k \in E \cup F\}, \quad (14)$$

to be the last event index not exceeding the i -th event, such that the buffer becomes either full or empty, and let

$$\chi_F(k) = \begin{cases} 1, & \text{if } k \in F, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

be the characteristic (indicator) function of set F .

Proposition 3.1 For every $i = 1, 2, \dots, N(\theta)$,

$$L'_i(\theta) = \begin{cases} \sum_{j=k(i)}^i \xi'_j(\theta) + \chi_F(k(i)) c'_{k(i)}(\theta) - c'_{i+1}(\theta), \\ 0, \end{cases} \quad \begin{array}{l} \text{if } i+1 \in F \\ \text{otherwise.} \end{array} \quad (16)$$

where $\xi_i(\theta) = [\alpha_i(\theta) - \beta_i(\theta)]\Delta t_i(\theta)$.

Proof. Suppose first that $i+1 \notin F$. Then the buffer does not become full in the interval $(t_i(\theta), t_{i+1}(\theta))$ for a small neighborhood of θ . Thus, $L_i(\theta) = 0$ and $L'_i(\theta) = 0$.

Suppose next that $i+1 \in F$. By Eq. (9),

$$L_i(\theta) = \xi_i(\theta) + x_i(\theta) - c_{i+1}(\theta),$$

which on differentiation yields

$$L'_i(\theta) = \xi'_i(\theta) + x'_i(\theta) - c'_{i+1}(\theta). \quad (17)$$

Consider the term $x'_i(\theta)$ above. Now, if $k(i) < i$, then for every $j = k(i) + 1, \dots, i$, the buffer is neither empty nor full in the interval $(t_{j-1}(\theta), t_j(\theta))$. Therefore, Eq. (1) implies

$$x_j(\theta) = x_{j-1}(\theta) + \xi_{j-1}(\theta),$$

which on differentiation yields

$$x'_j(\theta) = x'_{j-1}(\theta) + \xi'_{j-1}(\theta). \quad (18)$$

Combining Eqs. (17) and (18) results in the representation

$$L'_i(\theta) = \sum_{j=k(i)}^i \xi'_j(\theta) + x'_{k(i)}(\theta) - c'_{i+1}(\theta). \quad (19)$$

Finally, consider the term $x'_{k(i)}(\theta)$ above. By the definition of $k(i)$ in (14), either $k(i) \in E$ or $k(i) \in F$. In the first case, $x_{k(i)}(\theta) = 0$ and hence $x'_{k(i)}(\theta) = 0$, while in the second case, $x_{k(i)}(\theta) = c_{k(i)}(\theta)$ and hence $x'_{k(i)}(\theta) = c'_{k(i)}(\theta)$. In either case, Eq. (16) now follows from (19) and the definition of the characteristic function $\chi_F(k)$ in (15). \square

We next derive the derivative $L'(\theta; T)$ for several special cases.

Loss Volume as Function of Buffer Size

Let $c(\theta; t) = \theta$. However, let $\{\alpha(t)\}$ and $\{\beta(t)\}$ be independent of θ . Then, $\xi'_i(\theta) = 0$, while $c'_i(\theta) = 1$. An application of Eq. (16) reveals that for $i+1 \in F$, one has

$$L'_i(\theta) = \chi_F(k(i)) - 1. \quad (20)$$

Next, define a busy period of the buffer to be a maximal interval (period) during which the buffer is nonempty. For $i+1 \in F$, Eq. (20) then implies the following:

- If t_{i+1} is the time of the first event in its busy period at which the buffer is full, then $\chi_F(k(i)) = 0$, and hence $L'_i(\theta) = -1$.
- If t_{i+1} is *not* the time of the first event in its busy period at which the buffer is full, then $\chi_F(k(i)) = 1$, and hence $L'_i(\theta) = 0$.

It follows that $L'_i(\theta) = -1$ exactly once per busy period during which some loss was incurred. Letting $B(T)$ denote the number of busy periods during the interval $[0, T]$ during which some loss occurs, we have the result

$$L'(\theta; T) = -B(T). \quad (21)$$

Observe that the representation of $L'(\theta; T)$ in (21) is *distribution free*. We mention that this result has been derived in [6] in a different way.

Loss Volume as Function of Service Rate

Let θ be a parameter of the service rate, such that $\beta'_i(\theta) = 1$; in other words, the service rate increases uniformly in θ . However, let $\{\alpha(t)\}$ and $\{c(t)\}$ be independent of θ . Then

$$\begin{aligned} c'_i(\theta) &= 0 \\ \xi'_i(\theta) &= \frac{d}{d\theta}[\alpha_i - \beta_i(\theta)]\Delta t_i = -\Delta t_i. \end{aligned}$$

Eq. (16) then implies that

$$L'_i(\theta) = -\sum_{j=k(i)}^i \Delta t_j = t_{i+1} - t_{k(i)}, \quad i+1 \in F. \quad (22)$$

Note that the right-hand side above is just the elapsed time from the most recent time the buffer was either empty or full until the time t_{i+1} .

Next, we make two observations:

- $L'_{k(i)-1}(\theta) = 0$, $k(i) \in E$, by Eq. (16).

$$\bullet L'_{k(i)}(\theta) = - \sum_{j=k(i)}^{k(i)} \Delta t_j, \quad k(i) \in F, \text{ by Eq. (22).}$$

The above observations and Eq. (22) allow us to deduce that $L'(\theta; T)$ has the following structure. Let M be the number of busy periods in the interval $[0, T]$, during each of which the buffer becomes full at some point. Let τ_m be start time of the m -th such busy period, and let σ_m be the last event time in the m -th busy period at which the buffer is full. We then have the result

$$L'(\theta; T) = - \sum_{m=1}^M [\sigma_m - \tau_m]. \quad (23)$$

Observe that the representation of $L'(\theta; T)$ in (23) is *distribution free*.

Loss Volume as Function of Arrival Rate

Let θ be a parameter of the arrival rate, such that

$$\alpha'_i(\theta) = \begin{cases} 1, & \text{if } \alpha_i(\theta) > 0 \\ 0, & \text{if } \alpha_i(\theta) = 0 \end{cases}$$

In other words, the arrival rate increases uniformly in θ as long as there is an input flow. However, let $\{\beta(t)\}$ and $\{c(t)\}$ be independent of θ .

Next, we make the following observations:

- $c'_i(\theta) = 0$, since c_i does not depend on θ .
- If $\alpha_i(\theta) \neq 0$, then $\xi'_i(\theta) = \Delta t_i$.
- If $\alpha_i(\theta) = 0$, then $\xi'_i(\theta) = 0$.

Again, we consider the m -th busy period during which some loss occurs. Let σ_m and τ_m be as before, and let ℓ_m be the length of the interval, contained in the above m -th busy period, during which the inflow rate is 0, i.e., $\alpha(\theta; t) = 0$. An analysis along the lines of the previous case reveals that $L'(\theta; T)$ has a similar form except that the no-input intervals are accounted for. The resulting derivative is,

$$L'(\theta; T) = - \sum_{m=1}^M [\sigma_m - \tau_m - \ell_m]. \quad (24)$$

Again, the representation of $L'(\theta; T)$ in (24) is *distribution free*.

4 Conclusion

Formulas for the derivative of the loss volume as function of various parameters in a single-server CFM have been derived. They are simple to compute, distribution free, and provide unbiased derivative estimators. Therefore, they can be used in off-line design as well as in real-time flow control in high-speed packet networks. Future research will address extensions to networks with more general topology, as well as to various other performance metrics.

References

- [1] D. Anick, D. Mitra, and M.M. Sondhi, "Stochastic Theory of a Data-Handling System with Multiple Sources", *The Bell System Technical Journal*, Vol. 61, 1871–1894, 1982.
- [2] Y.C. Ho and X.R. Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*, Kluwer Academic Publishers, Boston, MA, 1991.
- [3] H. Kobayashi and Q. Ren, "A Mathematical Theory for Transient Analysis of Communications Networks", *IEICE Transactions on Communications*, Vol. E75-B, 1266–1276, 1992.
- [4] N. Miyoshi, "Sensitivity Estimation of the Cell-Delay in the leaky Bucket Traffic Filter with Stationary Gradual Input", *Proc. Intl. Workshop on Discrete Event Systems, WoDES'98*, 190-195, Cagliari, Italy, August 26-28, 1998.
- [5] Y. Wardi and B. Melamed, "IPA Gradient Estimation for the Loss Volume in Continuous Flow Models", *Proc. of the Hong Kong International Workshop on New Directions of Control and Manufacturing*, 30-33, Hong Kong, November 1994.
- [6] Y. Wardi and B. Melamed, "Modeling and Sensitivity Analysis of Continuous Flow Models", Manuscript, Georgia Tech, School of Electrical & Computer Engineering, and Rutgers University, Department of MSIS, 1999.