Bit dropping in ATM: A large deviations analysis

Adam Shwartz

Electrical Engineering
Technion—Israel Institute of Technology
Israel

Alan Weiss

AT&T Bell Laboratories
Murray Hill, NJ

A. Shwartz and A. Weiss,

Large deviations for performance analysis
queues, communication and computing,
Chapman-Hall, 1995.

ATM Conference, Paris, December 1995.

The basic Anick-Mitra-Sondhi model

i.i.d. on-off Markov fluid sources feed buffer

Capacity exceeds mean demand: $\frac{\lambda}{\lambda+\mu} < C$.

Buffer overflow rare but catastrophic

Control to prevent overflow.

Our goal (as number of sources increases)

- analyze effect of control
  on probability of rare event
- design control to be effective.

LARGE DEVIATIONS

- COMPUTE PROBABILITIES OF RARE EVENTS BY TRANSFORMING INTO VARIATIONAL PROBLEMS

LET $X_1, X_2, \ldots$ BE I.I.D. RANDOM VARIABLES, SET $S_n = n^{-1}\sum_1^n X_i$. DEFINE THE RATE FUNCTION

$$\ell(\alpha) \triangleq \sup_\theta \left\{\theta\alpha - \log \mathbb{E}e^{\theta X_1}\right\}. \tag{1}$$

Theorem 1 *Assume* $e^{\theta X_1} < \infty$ *for* θ *near* 0. *Then*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(S_n \in C\right) \le -\inf_{x\in C} \ell(x) \tag{2}$$

for every closed set C, *while for every open set* G,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(S_n \in C\right) \ge -\inf_{x\in G} \ell(x). \tag{3}$$

ℓ *is positive, convex, minimum* $(=0)$ *at* $\mathbb{E}X_1$.

THE SAME HOLDS FOR X_i IN $\mathbb{R}^d$, UNDER ADDITIONAL MILD ASSUMPTIONS.

$\boxed{\text{Large Deviations—Jump-Markov processes.}}$

Scaled Process z_n: increase local rates by n, decrease jump size by n. The question: compute

$$\mathbb{P}\left(\mathrm{z_n(t)\ follows\ a\ path\ r(t),\ 0 \le t \le T}\right)\ ? \tag{4}$$

Large Deviations for processes:

$$\mathbb{P}\left(\mathrm{z_n(t)\ near\ r(t)}\right) \approx \exp\left(-nI_0^T(r)\right) \tag{5}$$

$$I_0^T(r) \triangleq \int_0^T \ell(r(t), r'(t))\, dt. \tag{6}$$

An "event" is a set S of possible paths. So,

$$\mathbb{P}\left(\mathrm{z_n \in S}\right) \approx \exp\left[-\mathrm{n} \inf_{\mathrm{r \in S}} \mathrm{I_0^T(r)}\right] \tag{7}$$

Theorem 2 *Under technical assumptions,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\mathrm{z_n \in C}\right) \le -\inf_{\mathrm{r \in C}} \mathrm{I(r)}, \tag{8}$$

for every closed set C, while for every open set G,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\mathrm{z_n \in C}\right) \ge -\inf_{\mathrm{r \in G}} \mathrm{I(r)}. \tag{9}$$

If the event occurs, it does near the minimizing path.

The basic Anick-Mitra-Sondhi model

n i.i.d. Markov fluid sources, rate 1 each

Link capacity exceeds mean demand: $\frac{\lambda}{\lambda+\mu} < C$.

$$z_n(t) \triangleq \frac{1}{n} \cdot \text{number of ``on'' sources} \tag{10}$$

$$b_n(t) \triangleq \frac{1}{n} \cdot \text{buffer contents.} \tag{11}$$

Then z_n is Markov. Dynamics of buffer:

$$\frac{d}{dt} b_n(t) = \begin{cases} z_n(t) - C & \text{if } z_n(t) > C \text{ or } b_n(t) > 0 \\ 0 & \text{if } z_n(t) \leq C \text{ and } b_n(t) = 0 \end{cases} \tag{12}$$

$\boxed{\text{Filling the buffer to level } B}$

What is the probability?
How does this happen?

First make $z_n = C$ optimally.

Then (area under $z_n(t) - C$) $= B$ optimally.

First variational problem (q steady state):

$$\inf \left\{ I_0^T(r) : r(0) = q,\ r(T) = C,\ T > 0 \right\}. \qquad (13)$$

Free time! Second variational problem:

$$\inf \left\{ I_0^T(r) : r(0) = C,\ \int_0^T (r(t) - C)\, dt = B \right\}. \qquad (14)$$

Optimizer is a hyperbolic cosine: concave.

$\boxed{\text{From variational problem to probabilities}}$

What do we get from the solution
of the variational problem?

1. $\mathbb{P}\,(\text{buffer starts filling}) \approx \mathrm{e}^{-\mathrm{n}\mathrm{I}_\mathrm{c}^*}$,

where I_c^* solves level crossing problem.

2. If r_c^* is optimal, that is $I_c^* = I_c(r_c^*)$, then

$$\mathbb{P}\,(z_n \text{ near } r_c^* \,|\, \text{buffer starts filling}) \approx 1. \qquad (15)$$

3. $\mathbb{P}\,(\text{buffer overflows}) \approx \mathrm{e}^{-\mathrm{n}\mathrm{I}_\mathrm{b}^*}$,

where I_b^* solves overflow problem.

4. If r_b^* is optimal, that is $I_b^* = I_b(r_b^*)$, then

$$\mathbb{P}\,(z_n \text{ near } r_b^* \,|\, \text{buffer overflows}) \approx 1. \qquad (16)$$

That is, we obtain the probability and the path
of the unlikely event.

Asymptotics: small buffer

Approximate scaled rates.

$z =$ fraction of "on" sources.

$$\lambda(1-z) \to \lambda(1-C), \quad \mu z \to \mu C. \tag{17}$$

"Cost" $I^*(B)$ of filling B is

$$I^*(B) = c_2\sqrt{B} + O(B). \tag{18}$$

Asymptotics: large buffer

Most area obtained with $r(t) = \max_s r(s) = D$

where cost per buffer growth is minimal.

$$I^*(B) = c_3 B + c_4 + o(1). \tag{19}$$

$\boxed{\text{Control schemes}}$

Drop bits (lease significant) or packets (tagged)

When to start? when to stop?

$\boxed{\text{Rate based control:}}$

Reducing input when $z > D$ has no effect!

SO—DROP BITS according to BUFFER SIZE

Solution technique:

Split problem to pieces with smooth rates

Use principle of optimality on each piece

To match the pieces use

$\boxed{\text{Principle of smooth fit}}$

(Erdmann corner condition)

$\boxed{\text{Bit dropping schemes}}$

Trim above watermark $(\rho < 1,\ B > B_1)$

$$\frac{d}{dt}b_n(t) = \begin{cases} z_n(t) - C & \text{if } z_n(t) > C \text{ and } 0 < b_n(t) < B_1 \\ \rho z_n(t) - C & \text{if } b_n(t) > B_1 \\ 0 & \text{if } z_n(t) \le C \text{ and } b_n(t) = 0. \end{cases} \tag{20}$$

Principle of optimality: find cheapest way

$z_n(t) = C$, then

$b_n(t)$ exceeds B_1, and find resulting z_n

now make buffer B with input $\rho z_n(t)$.

Calculations—same as for the basic model

except possibly for values of some parameters!

$\boxed{\text{MORE BIT DROPPING SCHEMES}}$

$\boxed{\text{TRIM ALL PACKETS ABOVE WATERMARK}}$

ONCE WATERMARK EXCEEDED, TRIM

INCOMING AS WELL AS BUFFERED BITS

UNTIL BUFFER EMPTIES!

DYNAMICS: IN EACH BUSY PERIOD

IF WATERMARK EXCEEDED,

TRIM PACKETS IN BUFFER BY ρ, AND

$$\frac{d}{dt} b_n(t) =$$

$$\begin{cases} z_n(t) - C & \text{IF } b_n(t) > 0 \text{ AND } b_n(s) < B_1, \\ & \text{ALL } s < t \text{ IN CURRENT BUSY PERIOD} \\ \rho z_n(t) - C & \text{IF } b_n(t) > 0 \text{ AND } b_n(s) > B_1, \\ & \text{SOME } s < t \text{ IN CURRENT BUSY PERIOD} \\ 0 & \text{IF } z_n(t) \leq C \text{ AND } b_n(t) = 0 \end{cases}$$

NEED ADDITIONAL COORDINATE TO MAKE THIS MARKOV.

ANALYSIS IN SEGMENTS.

More bit dropping schemes

Two watermark

Start trimming if higher mark exceeded

Stop if went under lower watermark.

$$\frac{d}{dt}b_n(t) =$$

$$\begin{cases} z_n(t) - C & \text{if } b_n(t) > 0 \text{ and } b_n(s) < B_2, \\ & \text{all } s < t \text{ in current busy period} \\ z_n(t) - C & \text{if } 0 < b_n(t) < B_1 \\ \rho z_n(t) - C & \text{if } b_n(t) > B_1 \text{ and } b_n(s) > B_2, \\ & \text{some } s < t \text{ in current busy period} \\ 0 & \text{if } z_n(t) \leq C \text{ and } b_n(t) = 0 \end{cases} \tag{21}$$

Need additional coordinate

to make this Markov: trim status. Analysis in

segments.

Solution: usually numerical techniques

But can obtain explicit asymptotics.