On the compromise between burstiness and frequency of events

Alain Jean-Marie    Yvan Calas    Tigist Alemu

INRIA/LIRMM CNRS U. of Montpellier 2
ID-IMAG U. Grenoble
ajm@lirmm.fr    http://www.lirmm.fr/~ajm

Presented at
Performance 2005, 7 october 2005
with minor corrections
Contents

- Introduction
- Model 1
  - Experiments
- Model 2
  - Experiments
  - Analysis
- Application
  - Experiments
Consider a process of events with two types: **good** and **lost**. 

$n$ consecutive events are called a block. 

Time may be continuous, but the model will be in discrete-time and ignore actual time intervals between events.
Consider a process of events with two types: good and lost. 

\( n \) consecutive events are called a block.

Time may be continuous but the model will be in discrete-time and ignore actual time intervals between events.
The Problem (1)

Consider a process of events with two types: good and lost. 

$n$ consecutive events are called a block.

Time may be continuous but the model will be in discrete-time and ignore actual time intervals between events.
The Problem (2)

Metric of interest: given $h$ and $n$,

$$P( \text{the block is “lost”} ) = P( > h \text{ losses among } n \text{ events})$$

Usual objective: find the smallest $h$ (redundancy) such that

$$P( > h \text{ losses among } n + h \text{ events}) < \varepsilon.$$ 

Today’s objective: compare two situations

- same event loss probability
- different “burstiness” patterns
Motivation #1: Forward Error Correction

Forward error correction at the packet level: able to repair up to $h$ lost packets, using $h$ packets of redundancy.

$k=8$ information packets + $h=4$ redundancy packets

OK

LOST
Different queue management schemes at routers produce different loss patterns. Assuming the loss rate is the same: is it better

- to have losses regularly spaced,
- or have losses clustered?
Additional motivation

Reliability/real time systems:
- \( n \) tasks to be executed within a time frame
- each one may fail
- execute \( m = n + k \) of them
- “\( k\)-out-of-\( m \)”

Bandwidth reduction in a slotted network:
- frames of \( n \) slots \( \rightarrow \) frames of \( h \) slots
- no buffer
- probability of overflow?

Pedestrian crossing...
Several facts are well known:

- **Variability worsen things** (Folk result)
  \(\implies\) the situation with the “most regular arrivals” should be better

- The independence assumption is optimistic
  If the loss events are independent, the block loss probabilities are (much) smaller than if they are correlated
  \(\implies\) the situation with the “most independent arrivals” should be better

Investigate the issue with a focus on the bursts of losses.
Several facts are well known:

- **Variability worsen things** (Folk result)
  \[\Rightarrow\] the situation with the “most regular arrivals” should be better

- **The independence assumption is optimistic**
  If the loss events are independent, the block loss probabilities are (much) smaller than if they are correlated
  \[\Rightarrow\] the situation with the “most independent arrivals” should be better

Investigate the issue with a focus on the bursts of losses.
Several facts are well known:

- **Variability worsen things** (Folk result)
  \[\Rightarrow\] the situation with the “most regular arrivals” should be better

- **The independence assumption is optimistic**
  If the loss events are independent, the block loss probabilities are (much) smaller than if they are correlated
  \[\Rightarrow\] the situation with the “most independent arrivals” should be better

Investigate the issue with a focus on the **bursts of losses**.
**Assumption:** losses occur according to the state of a (two-state) Markov chain.

![Markov chain diagram]

The Gilbert model is described with the following assumptions:

- **Markov chain:** A (two-state) Markov chain is used to model the state transitions of the system.
- **Probabilities:** The transition probabilities are given by $a$ and $b$, where $1-a$ and $1-b$ represent the probabilities of staying in the current state.
- **Geometric distributions:** The time to return to the initial state follows a Geometric distribution with parameter $b$, and the time to enter the state 1 follows a Geometric distribution with parameter $a$.
The Gilbert model (2)

Gilbert as a Markov-Additive process:

\[ L_{m+1} = L_m + 1 \{ X_m = \bullet \} . \]

\[ E(z^{L_n}) = \pi_0 \ M(z)^n \ \mathbf{1} \]

\[ = (\pi_\bullet, \pi_\circlearrowleft) \times \left( \begin{array}{c} az \\ 1 - b \end{array} \right)^n \times \left( \begin{array}{c} 1 - a \\ b \end{array} \right) \]

where

\[ \pi_\bullet = \frac{1 - b}{2 - a - b} \quad \pi_\circlearrowleft = \frac{1 - a}{2 - a - b} . \]
Gilbert model (3)

Loss Run Length (LRL):

\[ \text{LRL} = \frac{1}{1 - a} \]

Good Run Length (GRL):

\[ \text{GRL} = \frac{1}{1 - b} \]

Stationary loss probability:

\[ p = \pi_0 = \frac{\text{LRL}}{\text{LRL} + \text{GRL}}. \]

Problem: with a fixed LRL (or \( a \)), the range of \( p \) is

\[ \left[ 0, \frac{\text{LRL}}{\text{LRL} + 1} \right). \]
Gilbert model (3)

Loss Run Length (LRL):

\[ \text{LRL} = \frac{1}{1 - a} \]

Good Run Length (GRL):

\[ \text{GRL} = \frac{1}{1 - b} \]

Stationary loss probability:

\[ p = \pi \cdot = \frac{\text{LRL}}{\text{LRL} + \text{GRL}}. \]

Problem: with a fixed LRL (or \(a\)), the range of \(p\) is

\[ \left[ 0, \frac{\text{LRL}}{\text{LRL} + 1} \right). \]
Skewed Gilbert Model

**Solution:** make Good Runs Geometrically distributed on \( \{0, 1, \ldots\} \) instead of \( \{1, 2, \ldots\} \).

\[ \begin{pmatrix} 1 - b(1-a) & b(1-a) \\ 1 - b & b \end{pmatrix}, \]

and

\[ \text{LRL} = \frac{1}{1-a}, \quad \text{GRL} = \frac{b}{1-b}, \quad p = \frac{1-b}{1-ab}. \]

Now the range of \( p \) is \([0, 1]\)!
Skewed Gilbert Model

**Solution**: make Good Runs Geometrically distributed on \{0, 1, \ldots\} instead of \{1, 2, \ldots\}.

\[ \Rightarrow \text{another Gilbert process with matrix:} \]
\[ \begin{pmatrix} 1 - b(1 - a) & b(1 - a) \\ 1 - b & b \end{pmatrix}, \]

and

\[ \text{LRL} = \frac{1}{1 - a} \quad \text{GRL} = \frac{b}{1 - b} \quad p = \frac{1 - b}{1 - ab}. \]

Now the range of \( p \) is \([0, 1]\)!
Skewed Gilbert Model

Solution: make Good Runs Geometrically distributed on \( \{0, 1, \ldots\} \) instead of \( \{1, 2, \ldots\} \).

\[ \implies \text{another Gilbert process with matrix:} \]

\[
\begin{pmatrix}
1 - b(1 - a) & b(1 - a) \\
1 - b & b
\end{pmatrix},
\]

and

\[
\text{LRL} = \frac{1}{1 - a} \quad \text{GRL} = \frac{b}{1 - b} \quad p = \frac{1 - b}{1 - ab}.
\]

Now the range of \( p \) is \([0, 1]\)!
Loss probability of a block of size $n = h + 16$, depending on $h$. 
Experiment: consider two cases 1 and 2.

- Fix the Loss Run lengths: \( LRL_1 < LRL_2 \) (\( a_1 < a_2 \)),
- fix a block length \( k \) and a “redundancy” quantity \( h \)
- vary the Loss Probability \( p \)
- plot the difference:

\[
\Delta_h(p) = P( \text{block saved in case 1} ) - P( \text{block saved in case 2} )
\]
Comparison experiments (2)

$h$ grows from 0 (left, red) to 13 (right, yellow).
Let $p_h$ be the value at which $\Delta_h(p_h) = 0$. 

![Graph showing the relationship between $h/15$ and critical probability $p_h$.](image-url)
Empirical finding: when $n$ is large,

$$x_h \sim \frac{h}{n - 1}.$$ 

How to prove it?

If the loss rate is $p = h/n$,

$$P(\leq h \text{ losses}) = [z^h] \frac{1}{1-z} \left( \frac{(1-c)z}{1-b} \frac{c z}{b} \right)^n$$

with

$$c = (1-b) \frac{n-h}{n}.$$ 

Work in progress...
**Simplification**: move to continuous time

![Diagram showing the transition from discrete to continuous time with marked intervals and events](image-url)
A Compound Poisson Model (1)

Process of loss:
- groups of losses occur according to a Poisson process with rate $\lambda$,
- groups have random sizes with identical distribution and mean $a$.

Global loss rate: $p = \lambda \times a$

Distribution of the number of losses:

$$\sum_{k} z^k P(k \text{ losses in } [0, t)) = e^{\lambda(A(z) - 1)}.$$
Comparison experiments (1)

Comparison of two cases:

- Small bursts case: losses of
  1 with proba 0.9,
  2 with proba 0.1
- Larger bursts case: losses of
  1 with proba 0.6,
  2 with proba 0.4
- Same average packet loss number \( x = p \times T \)

\[
\Delta_h(x) = P( \text{block saved with small bursts}) - P( \text{block saved with larger bursts})
\]
Comparison experiments (2)

Difference $\Delta_h(x)$ as the average number of losses $x$ grows

Again an empirical law

\[ x_h \sim h + C. \]
Analysis of limits

Analysis of extreme cases: consider the probability of success of a block

\[ P(N_T \leq h) = \sum_{n=0}^{h} \frac{x^n}{(E_A)^n n!} e^{-xT/E_A} P(A_1 + \ldots + A_n \leq h). \]

i/ Assume that

\[ \frac{P(A^{(1)} > h)}{E_A^{(1)}} < \frac{P(A^{(2)} > h)}{E_A^{(2)}}. \]

Then \( \Delta_h(x) > 0 \) when \( x \to 0 \).

ii/ Assume that \( m^{(1)} < m^{(2)} \). Then \( \Delta_h(x) < 0 \) when \( x \to \infty \).
Asymptotic Analysis (1)

Consider the quantity:

\[ d_h(y) = P( \leq h \text{ losses in } h + y \text{ time units}) . \]

Then we find:

\[
\begin{align*}
  d_h(y) &= \frac{1}{2} + \frac{1}{\sqrt{2\pi h}} \sqrt{\frac{\mu_1}{\mu_2}} \left( \frac{1}{2} + \frac{\mu_3}{2\mu_2} - y \right) + o(h^{-1/2}) ,
\end{align*}
\]

where \( \mu_1 = EA, \mu_2 = E(A^2), \mu_3 = E(A^3) . \)
Asymptotic Analysis (1)

Consider the quantity:

\[ d_h(y) = P( \leq h \text{ losses in } h + y \text{ time units}) . \]

Then we find:

\[
d_h(y) = \frac{1}{2} + \frac{1}{\sqrt{2\pi h}} \sqrt{\frac{\mu_1}{\mu_2}} \left( \frac{1}{2} + \frac{\mu_3}{2\mu_2} - y \right) + o(h^{-1/2}) ,
\]

where \( \mu_1 = EA, \mu_2 = E(A^2), \mu_3 = E(A^3) \).
Accordingly: for all real $y$, we have:

$$\Delta_h(h+y) = \frac{1}{\sqrt{2\pi h}} \left( C_0 - C_1 y \right) + o(h^{-1/2}),$$

where:

$$C_0 = \sqrt{\frac{\mu_1^{(1)}}{\mu_2^{(1)}} \left( \frac{1}{2} + \frac{\mu_3^{(1)}}{6\mu_2^{(1)}} \right)} - \sqrt{\frac{\mu_1^{(2)}}{\mu_2^{(2)}} \left( \frac{1}{2} + \frac{\mu_3^{(2)}}{6\mu_2^{(2)}} \right)},$$

and

$$C_1 = \sqrt{\frac{\mu_1^{(1)}}{\mu_2^{(1)}}} - \sqrt{\frac{\mu_1^{(2)}}{\mu_2^{(2)}}}.$$
Asymptotic analysis (3)

Finally, we have indeed:

$$\Delta_h(h + y_h) = 0 \implies y_h \sim \frac{C_1}{C_0},$$

and therefore

$$x_h \sim \frac{C_1}{C_0} + h.$$
Packet queues inside network routers are handled by a Queue Management scheme.

Two common ones:

- **Tail Drop**: Drops packets if and only if the buffer is full
  $$\implies$$ tends to produce bursts of losses

- **RED**: Drops packets at random preventively
  $$\implies$$ tends to produce isolated losses

Two loss patterns: which one works better with FEC?
Admitting that the smaller bursts (RED) work better when
\[ x \leq h + C \]
for some constant \( C \).

Equivalently, RED better if:

Small block
\[ k \leq \frac{1 - p}{p} h + \frac{C}{p} \]

Large redundant ratio
\[ \frac{h}{k} \geq \frac{1 - p}{1 - p k} - \frac{C}{1 - p} \]

Small loss rate
\[ p \leq \frac{h + C}{h + k} \]
Application of the model

Admitting that the smaller bursts (RED) work better when

\[ p(k + h) \leq h + C \]

for some constant \( C \).

Equivalently, RED better if:

- small block:
  \[ k \leq \frac{1 - p}{p} h + \frac{C}{p} \]

- large redund. ratio:
  \[ \frac{h}{k} \geq \frac{p}{1 - p} - \frac{C}{1 - p} \frac{1}{k} \]

- small loss rate:
  \[ p \leq \frac{h + C}{h + k} \]
Experimental setup

Simulations with the ns-2 program.

- Source of packets with the UDP protocol, 5-10% of the BW
- Background traffic of TCP flows, saturating the BW.

Statistics collected about Packet Loss Rate **Before Correction** and **after correction**.
Results of Simulations

**Loss rates**, $k = 16$ packets per block + $h = 2$ FEC packets.

RED does not always win...
There is a compromise between loss “burstiness” and loss rate. Assume blocks protected with $h = 1$ packet.
There is a compromise between loss “burstiness” and loss rate. Assume blocks protected with $h = 1$ packet.
There is a compromise between loss “burstiness” and loss rate. Assume blocks protected with $h = 1$ packet.