Large Deviations for Channels with Memory

Santhosh Kumar
Jean-Francois Chamberland
Henry Pfister

Electrical and Computer Engineering
Texas A&M University

September 30, 2011
Digital Landscape

Delay Tolerance and Rate Requirement

Context

- Packetized system supporting delay-sensitive traffic
- Point-to-point communication with partial feedback
Overview

Main Objective

▶ Analytic framework for delay sensitive systems
▶ Rigorous treatment of channels with memory

Highlights

▶ Focus on delay sensitive systems
▶ Random linear codes
▶ Large deviations

Potential Applications

▶ Optimal Code Rate
▶ Interface Selection
Main Result

- Large deviation principles on the delay and service rate for channels with memory.
Preview of Result

Main Result

- Large deviation principles on the delay and service rate for channels with memory

Why Large deviations?
Main Result

- Large deviation principles on the delay and service rate for channels with memory

Why Large deviations?

- Represents softer constraints on large buffers
Preview of Result

Main Result

- Large deviation principles on the \textit{delay} and \textit{service rate} for channels with memory

Why Large deviations?

- Represents softer constraints on large buffers
- Cumbersome to evaluate explicit probabilities
Main Result

- Large deviation principles on the delay and service rate for channels with memory

Why Large deviations?

- Represents softer constraints on large buffers
- Cumbersome to evaluate explicit probabilities
- Provides foundation for other concepts like effective capacity
Channel Abstraction - Finite State Bit Erasure Channel

- At each instant channel can be in one of $k$ states
- Channel transitions over time form a first-order Markov process
- In state $i$, transmitted bit is erased with probability $\varepsilon_i$
- Transition matrix

$$
\begin{bmatrix}
    p_{11} & \cdots & p_{1k} \\
    \vdots & \ddots & \vdots \\
    p_{k1} & \cdots & p_{kk}
\end{bmatrix}
$$

Evolution at bit level

$$
\begin{align*}
\text{State 1} & \quad 1 - \varepsilon_1 & b \\
& \quad \varepsilon_1 & e \\
\text{State 2} & \quad 1 - \varepsilon_2 & b \\
& \quad \varepsilon_2 & e \\
\text{State 3} & \quad 1 - \varepsilon_3 & b \\
& \quad \varepsilon_3 & e
\end{align*}
$$
Coding Scheme

Code Length $N$

- Information Bits $K$
- Parity Bits $N - K$

Code Generation
- Generate a binary random parity check matrix $H$ of size $(N - K) \times N$
- Codebook is null space of $H$

Code Properties
- Decoding failure probability (under ML) with $e$ erasures
  
  $$P_f(N - K, e) = 1 - \prod_{i=0}^{e-1} \left(1 - 2^{i-(N-K)}\right)$$
Distribution of Erasures induced by Channel

\[ \psi = \Pr(\text{Erasures}, C_{N+1} = j \mid C_1 = i) \]

\( \psi \) is the coefficient of \( x^{\text{Erasures}} \) in \( \begin{bmatrix} \begin{array}{c} P_{x_{11}}(1 - \epsilon_{1} + \epsilon_{1}x) \\ \vdots \\ P_{x_{1k}}(1 - \epsilon_{1} + \epsilon_{1}x) \\ \vdots \\ P_{x_{kk}}(1 - \epsilon_{k} + \epsilon_{k}x) \end{array} \end{bmatrix} \)
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

\[ \psi = \Pr(e \text{ Erasures, } C_{N+1} = j \mid C_1 = i) \]

Channel exit state  Channel beginning state
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

\[ \psi = \Pr(e \text{ Erasures, } C_{N+1} = j \mid C_1 = i) \]


- Channel exit state
- Channel beginning state
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

$$\psi = \Pr(\text{e Erasures, } C_{N+1} = j \mid C_1 = i)$$

Channel exit state  Channel beginning state
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

\[ \psi = \Pr(e \text{ Erasures, } C_{N+1} = j \mid C_1 = i) \]

Channel exit state \( C_{N+1} = j \)
Channel beginning state \( C_1 = i \)

\[ \begin{array}{cccc}
3 & 2 & 3 & 1 \\
\text{e} & \text{e} & \rightarrow (1 - \varepsilon_3)
\end{array} \]
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

\[ \psi = \Pr(e \text{ Erasures}, \quad C_{N+1} = j \quad | \quad C_1 = i) \]

\[ \begin{array}{c}
\text{Channel exit state} \\
\text{Channel beginning state}
\end{array} \]

\[ \rightarrow (1 - \varepsilon_3) \cdot p_{32} \varepsilon_2 \]
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

\[ \psi = \Pr(\text{e Erasures, } C_{N+1} = j \mid C_1 = i) \]

Channel exit state  Channel beginning state

- \( e \) Erasures, \( C_{N+1} = j \mid C_1 = i \)

- \( (1 - \epsilon_3) \cdot p_{32} \epsilon_2 \cdot p_{23} (1 - \epsilon_3) \cdot p_{31} \epsilon_1 \)
Distribution of Erasures induced by Channel

- Interested in conditional probabilities of type

\[ \psi = \Pr(e \text{ Erasures}, \quad C_{N+1} = j \mid C_1 = i) \]

\[ = \begin{cases} \text{Channel exit state} & \text{Channel beginning state} \end{cases} \]

\[ (1 - \varepsilon_3) \cdot p_{32} \varepsilon_2 \cdot p_{23} (1 - \varepsilon_3) \cdot p_{31} \varepsilon_1 \]

- \( \psi \) is the coefficient of \( x^e \) in \( [P^N_x]_{ij} \), where

\[ P_x = \begin{bmatrix}
  p_{11}(1 - \varepsilon_1 + \varepsilon_1 x) & \cdots & p_{1k}(1 - \varepsilon_1 + \varepsilon_1 x) \\
  \vdots & \ddots & \vdots \\
  p_{k1}(1 - \varepsilon_k + \varepsilon_k x) & \cdots & p_{kk}(1 - \varepsilon_k + \varepsilon_k x)
\end{bmatrix} \]
Packetized Model & Acknowledgments

- Acknowledgements of successful transmissions are available
- When transmission fails, data segments are re-transmitted
Hidden Markovian System
Hidden Markovian System

- System is Hidden Markov
Hidden Markovian System

- System is Hidden Markov
- Append Channel State
Hidden Markovian System

- System is Hidden Markov
- Append Channel State

Adding Channel State to Queue Length \((C_{sN+1}, q_s)\)
Hidden Markovian System

- System is Hidden Markov
- Append Channel State

Adding Channel State to Queue Length \((C_{sN+1}, q_s)\)

- The aggregate system is Markovian
Hidden Markovian System - System Evolution

Level Transitions

System Evolution

- Governed by Matrices: $M$ (Success) and $K$ (Failure)

$$M = \begin{bmatrix}
\mu_{11} & \mu_{12} & \mu_{13} \\
\mu_{21} & \mu_{22} & \mu_{23} \\
\mu_{31} & \mu_{32} & \mu_{33}
\end{bmatrix}$$

$$K = \begin{bmatrix}
\kappa_{11} & \kappa_{12} & \kappa_{13} \\
\kappa_{21} & \kappa_{22} & \kappa_{23} \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}$$
Delay & Service Rate - I

Delay

- $T_q$ is the time for delivery of data segment $q$
- Average Time:

$$Y_m = \frac{T_1 + \cdots + T_m}{m}$$

Service Rate

- $D_s$ (0 or 1) indicates the success of transmission $s$
- Average Service Rate:

$$Z_n = \frac{D_1 + \cdots + D_n}{n}$$
Delay & Service Rate - II

**Delay**

- $\tau$ - delay requirement
- Minimizing $\Pr(Y_m > \tau)$ is a service guarantee on delay

**Service Rate**

- $\gamma$ - service requirement
- Minimizing $\Pr(Z_n < \gamma)$ is a service guarantee on rate

**Basis for the large deviation principles**

- Look at the asymptotic decay rates (as $m, n \to \infty$) of $\Pr(Y_m > \tau)$ and $\Pr(Z_n < \gamma)$
- For large buffers large deviations characterization is appropriate
Large Deviations - A Brief Overview

► From WLLN, \( \Pr \left( \frac{X_1 + \cdots + X_n}{n} > \delta \right) \to 0 \), when \( \delta > E(X) \)

► Large deviations theory characterizes the exponential decay rate with which \( \Pr \to 0 \)

\[
I(\delta) = \lim_{n \to \infty} -\frac{1}{n} \log \Pr \left( \frac{X_1 + \cdots + X_n}{n} > \delta \right)
\]

Rate function

► When \( \{X_i\} \) are iid, Cramér’s Theorem characterizes the large deviations principle (LDP)

► When \( \{X_i\} \) are weakly dependent, invoke Gärtner-Ellis Theorem
Two Deviations

Theorem
$Y_m$ satisfies a large deviations principle with rate function $l_{del}(\cdot)$ where

\[ l_{del}(x) = \sup_{\lambda \in \mathbb{R}} \{ \lambda x - \log \rho((I - Ke^\lambda)^{-1}Me^\lambda) \} \]

Theorem
$Z_n$ satisfies a large deviations principle with rate function $l_{ser}(\cdot)$ where

\[ l_{ser}(x) = \sup_{\lambda \in \mathbb{R}} \{ \lambda x - \log \rho(K + Me^\lambda) \} \]

$\rho(\cdot)$ is the spectral radius of a Matrix
Relation between the deviations

- Set equivalence

\[
\{ T_1 + \cdots + T_m > n \} = \{ D_1 + \cdots + D_n < m \}
\]

- Relation

\[
x l_{del}(1/x) = l_{ser}(x)
\]

This implies the means (in the asymptotic sense) of \( \{D_s\} \) and \( \{T_q\} \) have an inverse relation.
Scaling for Commonality

Normalized Delay

\[ \tilde{Y}_m = \frac{N T_1 + \cdots + N T_{m/K}}{m} \]

\[ \frac{1}{K} I_{\text{del}} \left( \frac{K}{N} \tau \right) \]

Normalized Service Rate

\[ \tilde{Z}_n = \frac{K D_1 + \cdots + K D_{n/N}}{n} \]

\[ \frac{1}{N} I_{\text{ser}} \left( \frac{N}{K} \gamma \right) \]
GSM Example

System Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech frame length</td>
<td>228 bits</td>
</tr>
<tr>
<td>Physical layer</td>
<td>1 bit in 40µs</td>
</tr>
<tr>
<td>Delay Tolerance</td>
<td>40ms</td>
</tr>
</tbody>
</table>

- Requires 228 bits in 1000 channel uses
- Choose $\tau = \frac{1000}{228}$ for the example
Numerical Analysis

System Parameters

<table>
<thead>
<tr>
<th>Channel</th>
<th>Gilbert-Elliott Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erasure Probabilities</td>
<td>$\varepsilon_1 = 1, \varepsilon_2 = 0$</td>
</tr>
<tr>
<td>Channel Memory</td>
<td>0.9</td>
</tr>
<tr>
<td>Code length</td>
<td>N=114</td>
</tr>
</tbody>
</table>

Channel Diagram

- Average erasure rate is 20%
- $K$ is the parameter to be optimized
Numerical Analysis

Optimal Code Rate

- Depends heavily on the underlying traffic
- Very conservative as the service requirements are lowered
- The perspective of delay and service rate are related
Concluding Remarks

Main Contributions

▶ Provided a methodology for designing delay sensitive systems
▶ Rigorously modeled channels with memory

Possible Extensions

▶ Channel Estimation
▶ Analyze variable block lengths