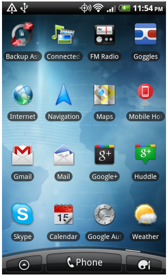
Information Theory, Queueing and Delay-Sensitive Traffic

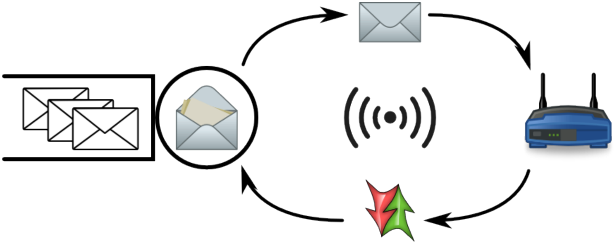
Jean-Francois Chamberland
Henry D. Pfister
Krishna Narayanan

Electrical and Computer Engineering
Texas A&M University

BIRS – Interactive Information Theory
Jan 17, 2012

Motivation and Target Infrastructures





Potential Scenarios

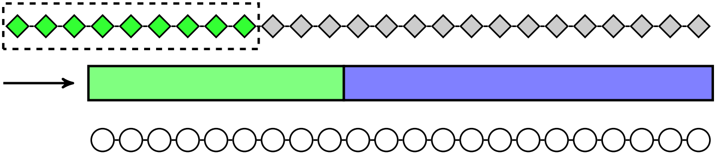
- Link layer
- Real-time applications
- Short blocks for delay

Structure

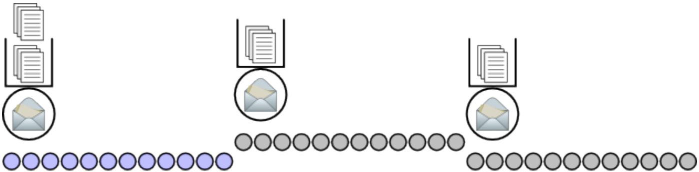
- Acknowledgements
- Partial state information
- Not *asymptopia*

Coding and Queueing

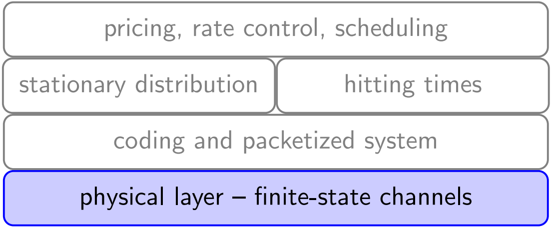
Block Encoding



Evolution of Transmit Buffer



Understanding the Performance of Queued Systems



Pertinent Features

- Short block codes over channels with memory
- Correlated service and queueing
- Interplay between system components

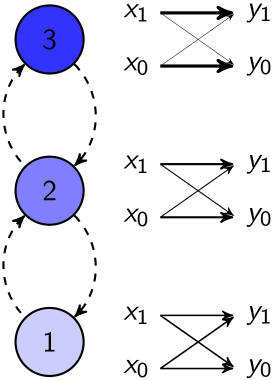
Finite-State Channels with Memory

Markov Channel Model

- Channel evolution forms Markov chain
- Probability of sequence $(c_1, c_2, \ldots c_t)$,

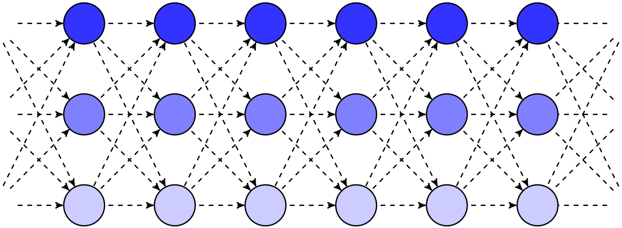
$$p_{C_1}(c_1) \prod p_{C_t|C_{t-1}}(c_t|c_{t-1})$$

- State determines channel quality
- Input-output relation is conditionally independent given state



Lifting over Time

Trellis of Possible Channel Realizations



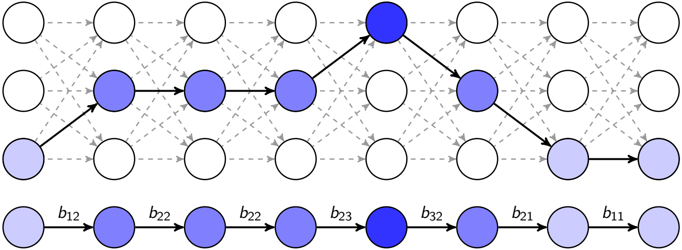
Transition Probability Matrix

- Channel memory
- Number of errors depends on empirical distribution of states

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

Channel Realization

Path through Trellis

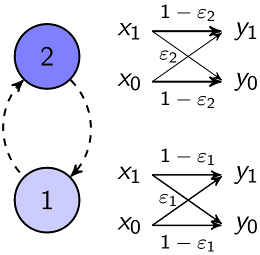


- Generating functions
- Efficient computation of empirical state distribution
- Accounting for all paths

$$\mathbf{B}(z_1, z_2, z_3) = \begin{bmatrix} b_{11}z_1 & b_{12}z_1 & b_{13}z_1 \\ b_{21}z_2 & b_{22}z_2 & b_{23}z_2 \\ b_{31}z_3 & b_{32}z_3 & b_{33}z_3 \end{bmatrix}$$

"Concrete Mathematics" by Graham, Knuth, Patashnik

Distributions (Gilbert-Elliott)



Generating Functions

- Summing weighted paths

$$\mathbf{B}^n(z_1, z_2) = \begin{bmatrix} b_{11}z_1 & b_{12}z_1 \\ b_{21}z_2 & b_{22}z_2 \end{bmatrix}^n$$

- Polynomial matrix

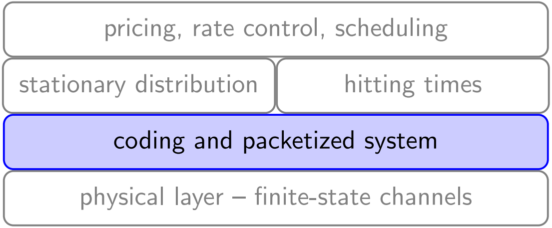
Empirical Probabilities

- Number of visits to each state

$$\Pr(N_1 = n_1, N_2 = n_2, C_{n+1} = j | C_1 = i) = [\![z_1^{n_1} z_2^{n_2}]\!] \left[\mathbf{B}^n(z_1, z_2)\right]_{i,j}$$

where $[\![\cdot]\!]$ is coefficient of monomial argument

Evolution of System at Link Layer

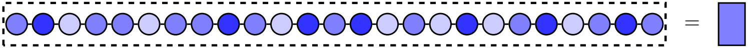


Major Points

- Intuition from empirical mutual information
- Random coding argument
- Preserve Markov property

Error Correcting Codes

Codeword and Successive Symbols



Analyzing Performance of Short Block Codes

- Intuition for channel state known at receiver
- Empirical average mutual information

$$\frac{1}{n} \log \left(\frac{\Pr(\mathbf{x}, \mathbf{y} | \mathbf{c})}{\Pr(\mathbf{x}) \Pr(\mathbf{y} | \mathbf{c})} \right) = \frac{1}{n} \sum_{t=1}^{n} \log \left(\frac{\Pr(x_t, y_t | c_t)}{\Pr(x_t) \Pr(y_t | c_t)} \right)$$

- Law large numbers and large deviations

Random Coding Argument

Binary Symmetric Channel

- Block length is n
- Number of codewords is m
- Maximum-likelihood decoder with ties as errors
- Probability of decoding failure

$$\sum_{e=0}^{n} \binom{n}{e} p^e (1-p)^{n-e} \left[1 - \left(1 - 2^{-n} \sum_{\tilde{e}=0}^{e} \binom{n}{\tilde{e}} \right)^{m-1} \right]$$

“Transmission of Information: A Statistical Theory of Communications” by R. M. Fano

Random Coding Argument

Gilbert-Elliott Channel

- State information at receiver
- Probability of decoding failure

$$\sum_{n_1+n_2=n} \Pr(N_1 = n_1, N_2 = n_2, C_{n+1} = j | C_1 = i)$$
$$\times \sum_{e_1=0}^{n_1} \binom{n_1}{e_1} \varepsilon_1^{e_1} (1-\varepsilon_1)^{n_1-e_1} \sum_{e_2=0}^{n_2} \binom{n_2}{e_2} \varepsilon_2^{e_2} (1-\varepsilon_2)^{n_2-e_2}$$
$$\times \left[1 - \left(1 - 2^{-n} \sum_{\gamma\tilde{e}_2+\tilde{e}_1 \leq \gamma e_2 + e_1} \binom{n_2}{\tilde{e}_2} \binom{n_1}{\tilde{e}_1} \right)^{m-1} \right]$$

where $\gamma = \frac{\log(\varepsilon_2/(1-\varepsilon_2))}{\log(\varepsilon_1/(1-\varepsilon_1))}$.

Horrible expression perhaps, but numerically efficient to compute

Alternate Possible Characterization

Model Structures

- Random linear codes over finite-state erasure channels with memory
- BCH codes over correlated error channels
- Gallager-type exponential upper bounds on error probability for finite-state channels with (increasing) memory

Key Components

- Probability of decoding failure

$$\Pr\left(\hat{w} \neq w, C_{n+1} = j \middle| C_1 = i\right)$$

- Keeping track of channel state

Time Required to Drain Buffer

pricing, rate control, scheduling

stationary distribution | **hitting times**

coding and packetized system

physical layer – finite-state channels

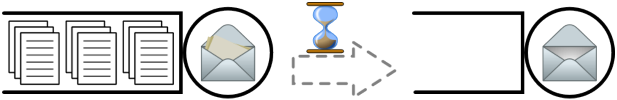
Delay and Service

- Distribution of completion time
- Leverage generating functions
- Time until service begin

Packetized Channel

Primary Goal

- Partition input bits into q segments of size k at time zero
- Find distribution (generating function) of first-passage time to empty queue

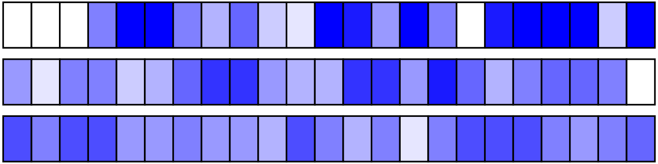


General Progression

- Identify evolution of packetized system from properties of underlying physical channel
- Apply analysis technique based on generating matrix

Channel Memory and Overall Performance

Memory Can Transcend Codeword Boundaries



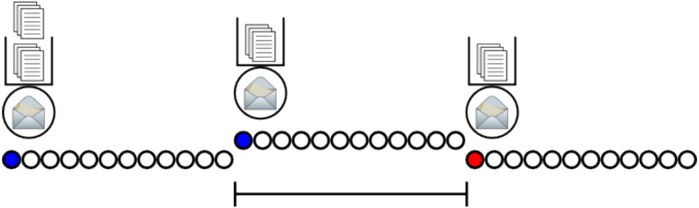
Low Correlation

- Concentration within codewords
- Average behavior
- Constant performance

Long Memory

- Dependence across codewords
- Highly predictable
- Opportunistic schemes

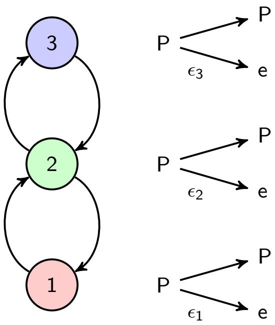
Markov Property and Aggregate Model



- Packetized system with Q_s segments
- Queue is hidden Markov
- Append channel state C_{sN+1} to queue length Q_s
- Aggregate process is Markov

Coded Channel and Packetized Behavior

Block Diagram



- Channel uncertainty produces packet erasures

Finite-State Service

- State i: packet erased with probability ϵ_i
- Decoding status observed

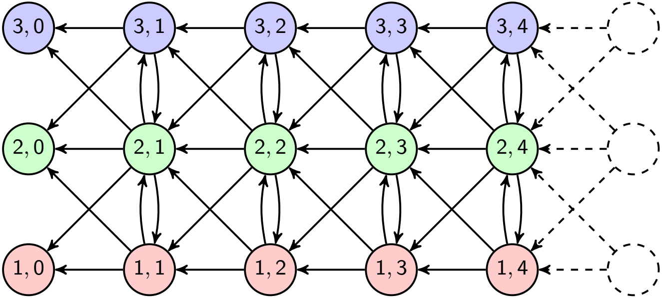
Evolution over Time

- Service is Markov
- Transition matrix

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

Lifting over Number of Segments in Queue

Adding Queue Length to Channel State

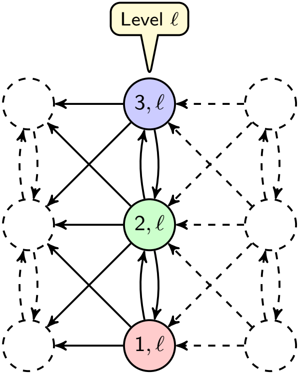


System Behavior

- Without arrivals, queue length can only decrease
- Adaptation is possible but not considered here

Level Decomposition

Level Structure



Quasi-Renewal Process

- Sum of conditionally independent random variables
- Implies product of generating matrices
- T_ℓ: Sojourn time at level ℓ
- H_0: Hitting time to level 0

$$\mathbf{G}_{H_0}(z) = \prod \mathbf{G}_{T_q}(z)$$

- Distribution of hitting time

Quasi-Renewal Structure

Take-Away Message

- Can efficiently calculate generating functions of hitting time to empty queue

$$G_H(z) = \mathrm{E}\left[z^H\right] \qquad M_H(s) = \mathrm{E}\left[e^{sH}\right] = G_H(e^s)$$

Performance Criteria of Choice

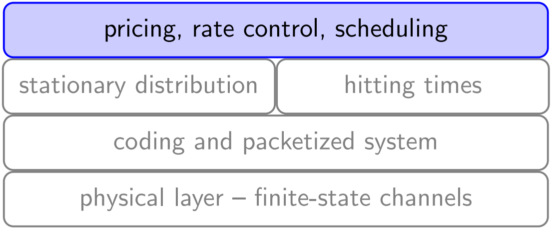
- Mean and compensated mean first-passage time

$$\mathrm{E}\left[H\right] = \frac{dM_H}{ds}(0) \qquad \mathrm{E}\left[He^{\lambda H}\right] = \frac{dM_H}{ds}(\lambda)$$

Desirable Properties of $\mathrm{E}\left[He^{\lambda H}\right]$

- Easy to compute & penalizes variations

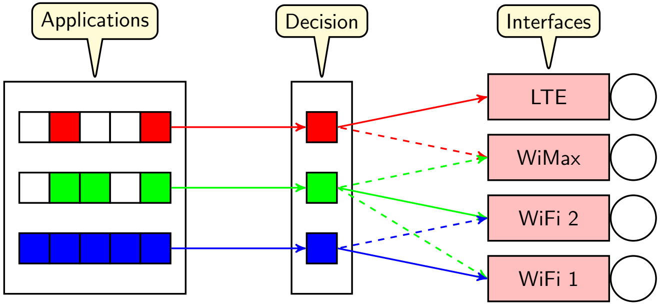
Multiple Interfaces and Heterogeneous Traffic



Objectives

- Route every packet to appropriate interface
- Account for needs, queues, channel profiles and conditions
- Adapt to traffic dynamics, activity levels and link quality

Interface Selection

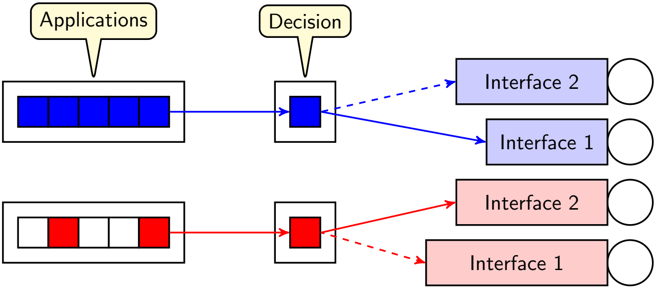


Objectives

- Requirements of heterogeneous traffic
- Soft matching of traffic to interfaces
- Sound modeling and rigorous analysis

Multiple Interfaces & Heterogenous Traffic

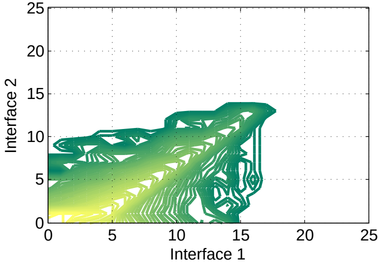
Decision Making Based on Compensated Queues

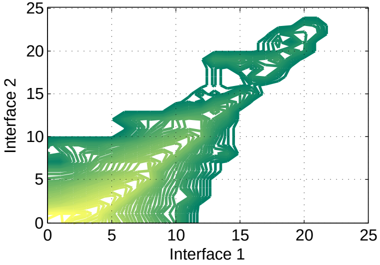


Same System under Different Lenses

- Application objectives: $\mathrm{E}[H]$ versus $\mathrm{E}[He^{\lambda H}]$ (delay-sensitive)
- Myopic sequential optimization – greedy algorithm

Applications Pick Favorable Conditions





Decision Based on $\mathrm{E}[He^{\lambda H}]$

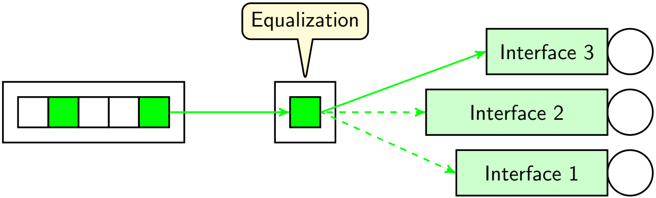
- Flow 1 is delay-sensitive
- Decision presents risk-adverse behavior
- Interface 1 is preferred

Decision Based on $\mathrm{E}[H]$

- Flow 2 tolerates variations
- Features better average but heavier tail
- Both interfaces leveraged

Multiple Interfaces & Heterogenous Traffic

Queue Selection Mechanism Works for Unitasking

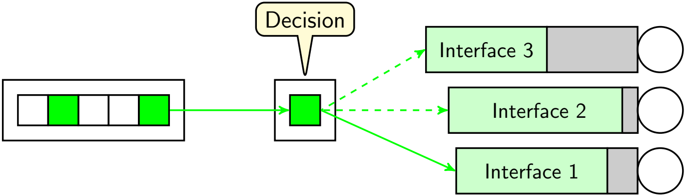


Challenges for Multitasking

- Bandwidth-hungry application dominates all queues
- Solution candidate: pricing or utility maximization framework
- Insights from effective bandwidth/capacity analysis
  1. Delay-sensitive traffic is costly
  2. Reliable channels are valuable

Pricing Framework

Advertised Price Influences Selection Process



Challenges for Multitasking

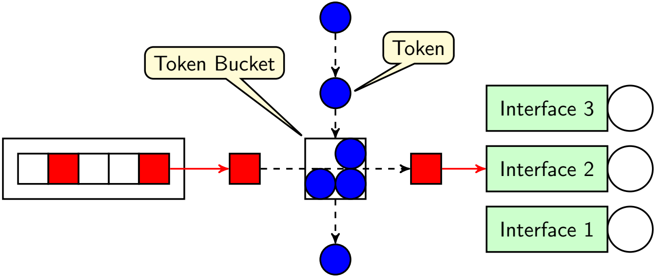
- Interface selected by minimizing cost

$$\text{interface}^* = \arg\min_i \mathrm{E}[H_i e_i^{\lambda H}] + P_i$$

where P_i is advertised price of sending packet on i

Rate Control and Optimization

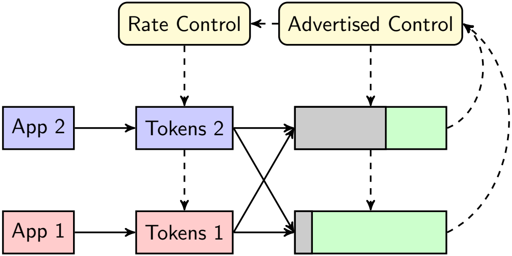
Rate Throttling through Token Bucket



- Rate control through token buckets
- Link price per packet for interface selection
- Compensated mean first-passage time
- Game theory and stochastic optimization

Overall System Architecture

Control Diagram



- Multilevel control policy
- Link prices set to optimize profit
- Possible interplay with network conditions
- Rate control (slowly) to maximize reward

Discussion and Concluding Remarks

Interface Selection Mechanism

- Rigorous analysis of queueing and coding
- Preserving Markov structure
- (weighted) Hitting time to head of queue
- Pricing framework for resource sharing
- Token bucket rate control with slow adaptation

Avenues of Future Research

- Pricing scheme for revenue extraction
- Link method to channel operation at symbol level
- Estimating system parameters
- Physical layer adaptation to load and scheduling algorithm