Rate control for communication networks

shadow prices, proportional fairness and stability

Author: FP Kelly, AK Maulloo and DKH Tan

Presented by Lei Ying

Introduction

- This paper analyses the stability and fairness of two classes of rate control algorithms for communication networks.
- The algorithms provide natural generalizations to large-scale networks of simple additive increase/multiplicative decrease schemes

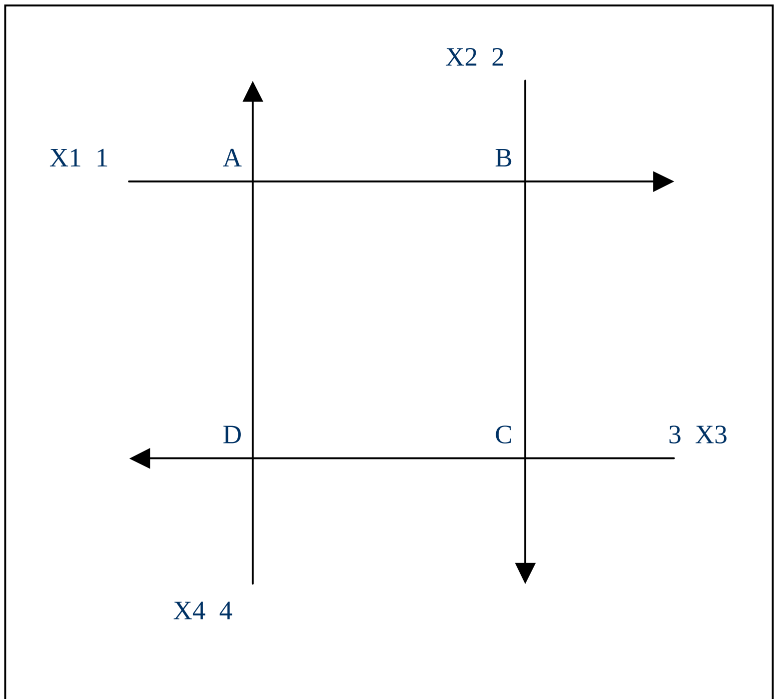
Model

Definition:

Consider a network with a set J of resources (abstraction of links), and let Cj be the finite capacity of resources j. Let a route r be a non-empty subset of J, and write R for the set of possible routes. Set Ajr=1 if $j \in r$ and Ajr=0 otherwise. This defines a 0-1 matrix A.

Associate a route r with a user, we allocate a rate x_r and a utility $U_r(x_r)$. Assuming $U_r(x_r)$ is increasing and strictly concave.

Model



- J={A, B, C, D}
- R={1, 2, 3, 4}
- 1 ={A, B}

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

1 2 3 4

- Users' rate: $x_1, x_2, x_{3,} x_4$

System Model

- SYSTEM(U, A, C)

$$\max \sum_{r \in R} U_r(x_r)$$

Subject to

$$Ax \leq C$$

over

$$x \geq 0$$

System Model

- This optimization problem is mathematically fairly tractable (with a strictly concave objective function and a convex feasible region), but it involves utilities U that are unlikely to be known by the network. We are thus led to consider two simpler problems.

User and Network Model

- Suppose that user r may choose an amount to pay per unit time, w_r, and receives in return a flow x_r proportional to w_r, say $x_r = w_r / \lambda_r$. We can get user model and network model:

User and Network Model

$$USER_r(U_r;\lambda_r):$$

$$\max U_r(\frac{w_r}{\lambda_r}) - w_r$$

over

$$w_r \geq 0$$

$$NETWORK(A,C;w):$$

$$\max \sum_{r \in R} w_r \log x_r$$

subject

$$Ax \leq C$$

over

$$x \geq 0$$

User and Network Model

- For User Model, the purpose is to maximize the user's gain (utility-payment).
- For Network, the purpose we choose Network Model as: $\max \sum_{r \in R} w_r \log x_r$ is to achieve fair. Define that the rates per unit charge of network is proportionally fair if x is feasible and for any other x^*

$$\sum_{r \in R} w_r \frac{x_r^* - x_r}{x_r} \leq 0 \Rightarrow \sum_{r \in R} w_r \frac{\delta x_r}{x_r} \leq 0$$

User and Network Model

- It means the smaller flows and more generous users will be favored.
- If x is a optimum solution to NETWORK, we have

$$(x_r, r \in R) \rightarrow (x_r + \delta x_r, r \in R) \Rightarrow \sum_{r \in R} w_r \frac{\delta x_r}{x_r} \leq 0$$

It is same as the above condition.

So when we maximize NETWORK, it is also insure that the rates per unit charge of network is proportionally fair.

User and Network Model

- It is proved in Kelly’s paper “Charging and rate control for elastic traffic” that there always $\exists \lambda$ w x , satisfying $w_r = \lambda_r x_r$ such that , w_r solves USERr and vector x solves NETWORK. Further x is the unique solution to SYSTEM and rates per unit charge of the NETWORK is proportionally fair.

User and Network Model

- Under the decomposition of the problem SYSTEM into the problems NETWORK and USERr, the utility function Ur is not required by the network.

  So suppose network knows w_r , we design algorithms to solve the NETWORK problem.

Primal Algorithm

$$\frac{d}{dt}x_r(t) = k\left(w_r - x_r(t)\sum_{j\in r}\mu_j(t)\right) \qquad (1)$$

Where

$$\mu_j = p_j\left(\sum_{s=j\in s} x_s(t)\right) \qquad (2)$$

$p_j(y)$ means that when the total flow of j is y, the unit flow charge is p_j

Primal Algorithm

- Define

$$u(x) = \sum_{r \in R} w_r \log x_r - \sum_{j \in J} \int_0^{\sum_{s: j \in s} x_s} p_j(y)\,dy$$

- Theorem 1: The strictly concave function $u(x)$ is a Lyapunov function for the system of differential equations (1)-(2). The unique value x maximizing $u(x)$ is a stable point of the system, to which all trajectories converge.

Primal Algorithm

- Proof: $u(x)$ is strictly concave on $x \geq 0$ with an interior maximum.

$$\frac{\partial}{\partial x_r} u(x) = \frac{\omega_r}{x_r} - \sum_{j \in r} p_j \left(\sum_{s: j \in s} x_s \right)$$

$$\frac{d}{dt} u(x(t)) = k \sum_{r \in R} \frac{1}{x_r(t)} \left(w_r - x_r(t) \sum_{j \in r} p_j \left(\sum_{s: s \in j} x_x(t) \right) \right)^2$$

establishing that $u(x)$ is strictly increasing with t, unless the unique x maximizing $u(x)$. The function $u(x)$ is thus a Lyapunov function for system above.

Primal Algorithm

- Define the continuous function:

$$p_j(y) = \frac{\left(y - C_j + \varepsilon\right)^+}{\varepsilon^2} = \mu_j$$

$\varepsilon \to 0 \Rightarrow$ maximization of the lyapunov function approximates arbitrarily closely the primal problem NETWORK.

Rate of convergence

- Let $x_r(t) = x_r + x_r^{\frac{1}{2}} y_r(t)$, linearizing

$$\frac{d}{dt} y(t) = -k\left(WX^{-1} + X^{\frac{1}{2}} A^T P' A X^{\frac{1}{2}} \right) y(t)$$

$$X = diag(x_r, r \in R) \quad W = diag(w_r, r \in R) \quad P' = diag(P_j', j \in J)$$

$$\Gamma^T \Phi \Gamma = WX^{-1} + X^{\frac{1}{2}} A^T P' A X^{\frac{1}{2}}$$

$$\Gamma^T \Gamma = \mathrm{I} \qquad \Phi = diag(\Phi_r, r \in R)$$

we get $\frac{d}{dt} y(t) = -K \Gamma^T \Phi \Gamma\, y(t)$

Rate of convergence

- Rate of convergence is determined by the smallest eigenvalue, Φ_r. And speed of convergence increases both with gain parameter K and P'.

$K\uparrow$	Speed of convergence$\uparrow$
$P^{'}\uparrow$	Speed of convergence$\uparrow$

Table 1

Stochastic analysis

- We consider a stochastic perturbation of linearized equation. Let

$$dy(t) = -\kappa\left(\Gamma^T \Phi \Gamma y(t)dt + FdB(t)\right)$$

B(t) is a collection of independent standard Brownian motions.

The stationary solution is:

$$y(t) = -k \int_{-\infty}^{t} e^{-k(t-\tau)\Gamma^T \Phi \Gamma} FdB(\tau)$$

It has a multivariate normal distribution

$$y(t) \sim N(0, \Sigma)$$

Stochastic analysis

- So if we define

$$[\Gamma F;\Phi]_{rs} = \left[\int_{-\infty}^{0} e^{\tau\Phi}\Phi F F^{T}\Gamma^{T} e^{\tau\Phi}\, d\tau\right]_{rs} = \frac{[\Gamma F F^{T}\Gamma^{T}]_{rs}}{\Phi_{r} + \Phi_{s}}$$

We get

$$\Sigma = k\Gamma^{T}[\Gamma F;\Phi]\Gamma$$

$K \uparrow$	$\Sigma \uparrow$
$P' \uparrow$	$\Sigma \downarrow$

Table 2

As P’ increases, not only is convergence faster, but also the spread at equilibrium decreases.

Various stochastic Model

- We have various sources of randomness that may lead to different covariance structure.
- Congestion indication with join feedback:

$$dx_r(t) = k\left(\omega_r dt - x_r(t)\sum_{j\in r}\varepsilon_j dN_j\left(\varepsilon_j^{-1}\int_0^t \mu_j(\tau)d\tau\right)\right)$$

Various stochastic Model

- Resource j generates feedback signal indicating congestion as a time-dependent Possion process at rate $\varepsilon_j^{-1}\mu_j(t)$

  suppose that when a feedback signal is generated, it is send to each user r whose route passes through resource j, and the rate x_r reduces $\kappa\varepsilon x_r(t)$

Various stochastic Model

- We can get the corresponding Brownian version of the linearized system let

$$F_{rj} = \varepsilon_j^{\frac{1}{2}} \mu_j^{\frac{1}{2}} A_{jr} x_r^{\frac{1}{2}}$$

$$FF^T = X^{\frac{1}{2}} A^T EPAX^{\frac{1}{2}}$$

Various stochastic Model

- Congestion indication with individual feedback:

$$dx_r(t) = k\left(\omega_r dt - \sum_{j\in r}\varepsilon_j dN_{jr}\left(\varepsilon_j^{-1}\int_0^t x_r(\tau)\mu_j(\tau)d\tau\right)\right)$$

Feedback signals from resource j to user r arise at rate $\varepsilon_j^{-1}x_r(t)\mu_j(t)$. And user r reacts to such a feedback signal by reducing x_r with $\kappa\varepsilon$.

We get

$$F_{r,(j,s)} = \varepsilon_j^{\frac{1}{2}}\mu_j^{\frac{1}{2}}A_{jr}\mathrm{I}[r=s]$$

Time Lags

Consider next the lagged, discrete time system

$$x_r[t+1] = x_r[t] + k\left(w_r - x_r[t]\sum_{j\in r}\mu_j[t-d(j,r)]\right)$$

$$\mu_j[t] = p_j\left(\sum_{s:j\in s} x_s[t-d(j,s)]\right)$$

$d(j,s)$ is the delay between user s changing its rate and the altered flow reaching resource j.

$d(j,r)$ is the delay between resource j generating a feedback signal and user r receiving it

Discrete System with Time Lags

Theorem 2:

The vector x maximizing the strictly concave function $u(x)$ is the unique equilibrium point of this system.

- Proof: The vector x is an equilibrium point if and only if it solves

$$w_r = x_r \sum_{j \in r} p_j \left(\sum_{s: j \in s} x_s \right)$$

It is precisely the stationarity condition of system (1)-(2)

Discrete System with Time Lags

- If we let $\mu_j = p_j\left(\sum_{s:j\in s} x_s\right)$, and suppose p_j is differentiable at point x_r. Let $x_r[t] = x_r + x_r^{1/2} y_r[t]$

Then, linearizing the above system about x.

If we define

$$(L[d])_{rs} = \sum_j p_j{}' A_{jr} A_{js} x_r^{\frac{1}{2}} x_s^{\frac{1}{2}} \mathrm{I}[d(j,r)+d(j,s)=d]$$

$$\sum_{d=0}^{D} L[d] = X^{\frac{1}{2}} A^T P' A X^{\frac{1}{2}}$$

Discrete System with Time Lags

$$L = \begin{pmatrix} \mathrm{I} - k\left(WX^{-1} + L[0]\right) & -kL[1] & -kL[2] & \ldots & -kL[D] \\ \mathrm{I} & 0 & 0 & \ldots & 0 \\ 0 & \mathrm{I} & 0 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

$$D = \max_{j,r,s}\{d(j,r) + d(j,s)\}$$

Discrete System with Time Lags

$$\begin{pmatrix} y[t+1] \\ y[t] \\ \cdots \\ y[t-D+1] \end{pmatrix} = L \begin{pmatrix} y[t] \\ y[t-1] \\ \cdots \\ y[t-D] \end{pmatrix}$$

- The equilibrium point x of the system is stable if and only if the spectral radius of the matrix L is less than unity.

Discrete System with Time Lags

$K\uparrow$	The system goes unstable
$P'\uparrow$	The system goes unstable

Table 3

Dual System

- The Lagrangian for the problem NETWORK is

$$L(x,z;\mu) = \sum_{r\in R} w_r \log x_r + \mu^T (C - Ax - z)$$

where $z \geq 0$ is a vector of slack variable and μ is a vector of lagrange multipliers (or shadow price). Then

$$\frac{\partial L}{\partial x_r} = \frac{w_r}{x_r} - \sum_{j\in J} A_{jr}\mu_{jr}$$

So the unique optimum to the primal problem is given by

$$x_r = \frac{w_r}{\sum_{j\in J} A_{jr}\mu_{jr}}$$

Dual System

Where $x_r \ \mu_j$ solve

$$\mu \geq 0, Ax \leq C, \mu^T (C - AX) = 0$$

Maximizing the Lagrangian over x and z gives

$$L(x, z; \mu) = \sum_{r \in R} w_r \log\left(\frac{w_r}{\sum_{j \in J} A_{jr} \mu_{jr}} \right) + \mu^T C - \mu^T Ax$$

$$= \sum_{r \in R} w_r \log(w_r) - \sum_{r \in R} w_r \log\left(\sum_{j \in J} A_{jr} \mu_{jr} \right) + \sum_{j \in J} \mu_j C_j - \sum_{r \in R} w_r$$

Dual System

From it, removing the terms not dependent on the shadow price μ_j we can derive the DUAL:

$$\max \sum_{r \in R} w_r \log\left(\sum_{j \in J} \mu_j \right) - \sum_{j \in J} \mu_j C_j$$

Dual System

- Also we have a dual algorithm is:

$$\frac{d}{dt}\mu_j(t) = k\left(\sum_{r:j\in r} x_r(t) - q_j(\mu_j(t))\right) \qquad (3)$$

$$x_r(t) = \frac{w_r}{\sum_{k\in r}\mu_k(t)} \qquad (4)$$

$q_j(\eta)$ means the flow through resource j is $q_j(\eta)$, it generates a price η at resource j

Dual System

- If we define

$$\upsilon(\mu) = \sum_{r \in R} w_r \log\left(\sum_{j \in r} \mu_j\right) - \sum_{j \in J} \int_0^{\mu_j} q_j(\eta)\, d\eta$$

- Theorem 3: The strictly concave function υ is a Lyapunov function for the system .The unique value μ maximizing υ is a stable point of the system, to which all trajectories converge.

Dual System

- We also have rate of convergence:

$$\frac{d}{dt}\xi(t) = -k\Theta^T\Psi\Theta\xi(t)$$

$K\uparrow$	Speed of convergence $\uparrow$
$Q^{'}\uparrow$	Speed of convergence $\uparrow$

Table 4

Dual System

- Stochastic analysis: $d\xi(t) = -k\left(\Theta^T \Psi \Theta \xi(t) dt - G dB(t)\right)$

let $[\Theta G;\Psi]_{jk} = \dfrac{\left[\Theta G G^T \Theta^T\right]_{jk}}{\psi_j + \psi_k}$

then the covariance matrix is $\Sigma = k\Theta^T[\Theta G;\Psi]\Theta$

$K \uparrow$	$\Sigma \uparrow$
$Q' \uparrow$	$\Sigma \downarrow$

Table 5

Dual System

- Discrete system with time lags:

$$\mu_j[t+1] = \mu_j[t] + k\left(\sum_{r:j\in r} x_r[t - d(j,r)] - q_j(\mu_j[t])\right)$$

$$x_r[t] = \frac{w_r}{\sum_{k\in r} \mu_k[t - d(k,r)]}$$

Dual System

$$\begin{pmatrix} \xi[t+1] \\ \xi[t] \\ \dots \\ \xi[t-D+1] \end{pmatrix} = M \begin{pmatrix} \xi[t] \\ \xi[t-1] \\ \dots \\ \xi[t-D] \end{pmatrix}$$

$$(M[d])_{jk} = \sum_r x_r A_{jr} A_{kr} \mathrm{I}[d(j,r) + d(k,r) = d]$$

Dual System

$$M = \begin{pmatrix} I - k\left(M[0] + Q'\right) & -kM[1] & -kM[2] & \dots & -kM[D] \\ I & 0 & 0 & \dots & 0 \\ 0 & I & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

$K \uparrow$	The system goes unstable
$Q' \uparrow$	The system goes unstable

Table 6

Four nodes network

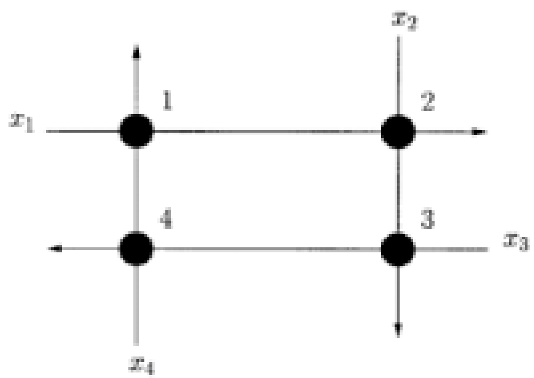


Figure 1 A four node network.

Suppose that the total load y on a resource takes the form of a Poisson stream of cells at rate y/ε.

Suppose that the time-axis is divided into non-overlapping slot $\tau\varepsilon$, and that a feed back signal is generated for a slot if the total number of cells arriving in that slot exceeds a threshold N.

Four nodes network

- Joint feedback:suppose that when a feedback signal is generated, it is send to each user r whose route passes through resource j, and the rate x_r reduces $\kappa\varepsilon x_r(t)$
- Individual feedback Suppose that when a feedback signal is generated at resource j, it is directed at a random route r with probability $\frac{x_r}{y}$, and rate x_r reduced $\kappa\varepsilon$

Four nodes network

- Consider the network illustrated above, where $|J| = |R| = 4$
  let $w_r = 0.0002$,and choose $N = 128$
  $\tau = 50$ so that $x_r = 1.0$ and $\mu_j = 0.0001$

Four nodes network

$$\Sigma_1 = \kappa\varepsilon 10^{-2}\begin{pmatrix} 2.4 & 0.8 & -0.8 & 0.8 \\ 0.8 & 2.4 & 0.8 & -0.8 \\ -0.8 & 0.8 & 2.4 & 0.8 \\ 0.8 & -0.8 & 0.8 & 2.4 \end{pmatrix}$$

$$\Sigma_2 = \kappa\varepsilon 10^{-1}\begin{pmatrix} 1.5 & -1.2 & 1.1 & -1.2 \\ -1.2 & 1.5 & -1.2 & 1.1 \\ 1.1 & -1.2 & 1.5 & -1.2 \\ -1.2 & 1.1 & -1.2 & 1.5 \end{pmatrix}$$

Four nodes network

- Compare the magnitudes and structures of the two matrices.
- We can see Σ_2 is larger in magnitude, since with individual feedback there are additional sources of variation in the random choice of which rate is to be reduced.

Four nodes network

- Note also that rates on routes sharing a node are positively correlated in joint feed back, because congestion indication at a node causes both routes through that node to decease. However, for individual feedback, routes sharing a node are negatively correlated. A decrease in the flow on a route will allow increase on routes sharing a common node with it.

A random network

- We consider a network where the elements of the matrix A are independent random variables, each taking the value 1 with probability p and the value 0 otherwise.
- Let $w_r = \sum_j A(j,r)$ and $q_j(\eta) = \eta \sum_s A(j,s)$ so the unique stable point for the system (3)-(4) is $x_r = 1$ and $\mu_r = 1$

A random network

- Consider the system:

$$\mu_j[t+1] = \mu_j[t] + \kappa\left(\sum_{r:j\in r} N_r[t-d(j,r)] - q_j(\mu_j[t])\right)$$

- Where $N_r[t]$ are collection of independent Poisson random variables, and has mean $x_r[t]$. Choose five random routes and resources for the following parameter: J=100, R=1000, p=0.1, k=0.005.

A random network

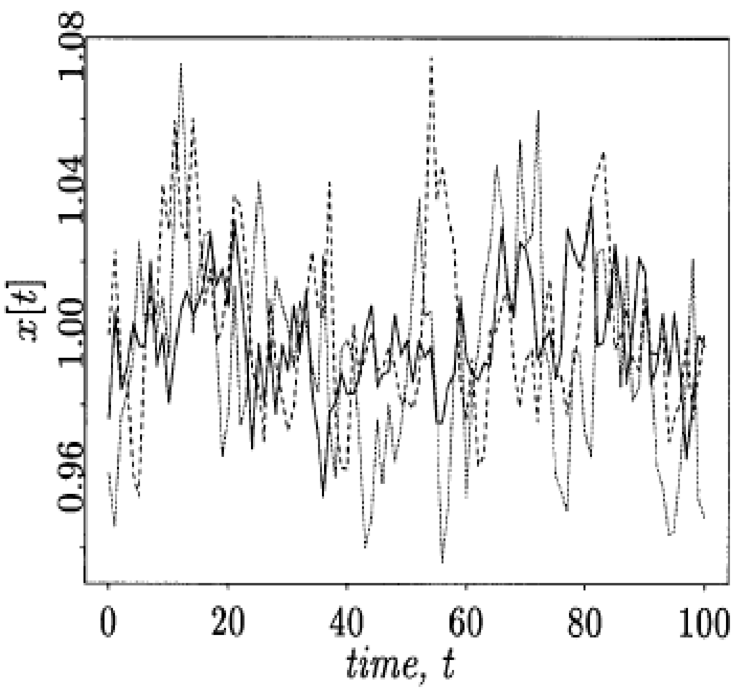


Figure 2 Rates on three randomly chosen routes.

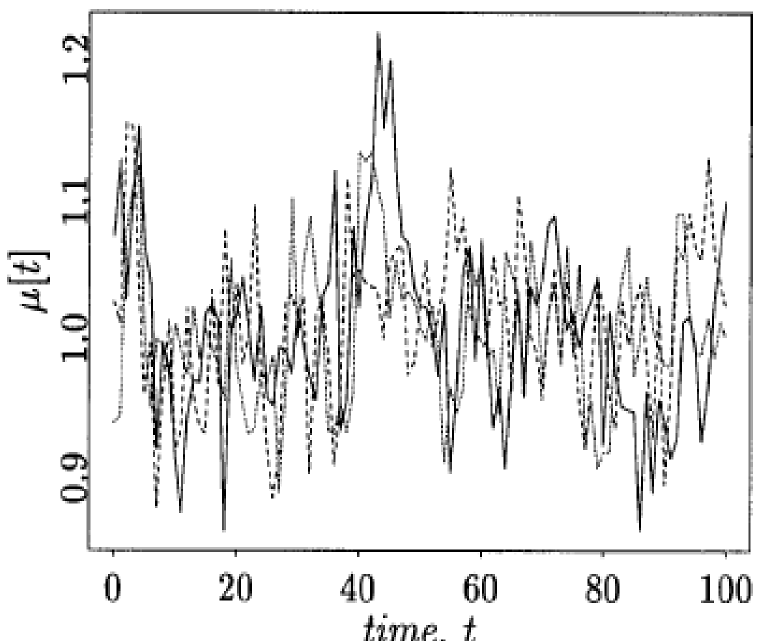


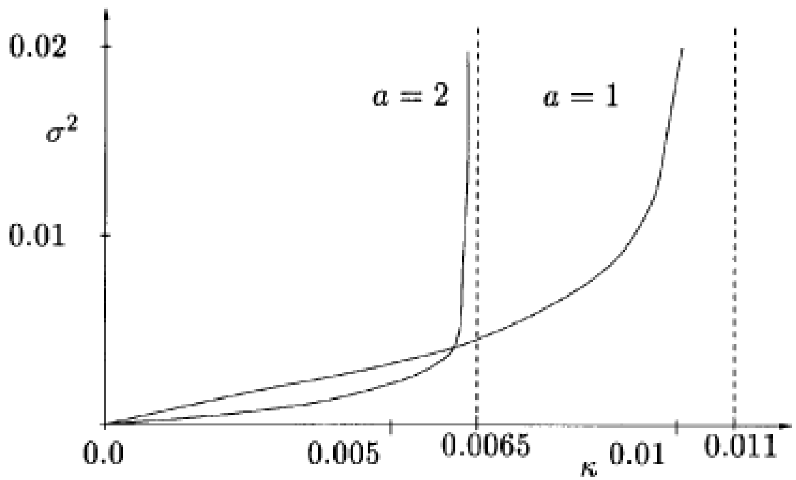
Figure 3 Shadow prices for three randomly chosen resources.

A random network

- From the above figures, we can see that for this example rates oscillate within a narrower band than shadow prices, and both are relatively well controlled.
- Finally, let us consider briefly the effect of more general choices for q_j , let

$$q_j = (a(\eta - 1) + 1)\sum_s A(j, s)$$

A random network



The case a=1 is discussed so far; the case a=2, corresponds to a doubling of the matrix of derivatives Q'. As predicted, increasing a is to reduce variability, but also to lower the critical value of k at which the system becomes unstable

User adaptation

- Suppose user r is able to monitor its rate $x_r(t)$ and vary $w_r(t)$ to track the optimum of USERr. so we will always have $w_r = x_r U_r'(x_r)$

  For the primal algorithm:

$$\frac{d}{dt}x_r(t) = k\left(w_r(t) - x_r(t)\sum_{j\in r} w_j(t) \right)$$

User adaptation

- The revised differential equation

$$u(x) = \sum_{r \in R} U_r(x_r) - \sum_{j \in J} \int_0^{\sum_{s: j \in s} x_s} p_j(y)\,dy$$

- Provides a lyapunov function for the revised system, and the unique value maximizing u(x) is a stable point of the system, to which all trajectories converge.

User adaptation

- Linearization may again be used to investigate behavior near the stable point:

$$\frac{d}{dt}y(t) = -kX^{\frac{1}{2}}\left(A^T P' A - U''\right)X^{\frac{1}{2}}y(t)$$

$$U'' = diag\left(U_r''(x_r), r \in R\right)$$

End

Thanks!