Scheduling and Control of Mobile Communications Networks with Randomly Time Varying Channels by Stability Methods

Harold J. Kushner
Applied Mathematics Dept., Brown University
Providence RI 02912 USA
hjk@dam.brown.edu

Abstract— Consider a communications network consisting of mobiles, some of which can serve as a receiver and/or transmitter in a multihop path. There are random external data processes, each destined for some destinations. At each mobile the data is queued according to the source-destination pair until transmitted. The capacities of the connecting channels are randomly varying. Time is divided into small scheduling intervals. At the beginning of the intervals, the channels are estimated via pilot signals and this information is used for the scheduling decisions during the interval, concerning the allocation of transmission power and/or time, bandwidth, and perhaps antennas, to the various queues in a queue and channel-state dependent way, to assure stability. Lost packets might or might not have to be retransmitted. General networks are covered, conditions used in previous works are weakened, and the distributions of the input file lengths can be heavy tailed. The resulting controls are readily implementable. The choice of Liapunov function allows a range of tradeoffs between current rates and queue lengths, under very weak conditions. Because of the non-Markovian nature of the problem, we use the perturbed Stochastic Liapunov function method, which is designed for such problems. Extensions concerning acknowledgments, multicasting, non-unique routes, and others, are available.

Index terms: Scheduling in stochastic networks, randomly-varying link capacities, mobile networks, stochastic stability, stability of networks with randomly varying links, routing in ad-hoc networks, perturbed stochastic Liapunov functions, heavy tailed distributions.

I. Introduction

Consider a network of M mobiles or nodes. There are S external sources of bursty data processes, each source having a unique entry and destination node in the system. In this short paper, the routing is a priori fixed for each of the source-destination pairs. This simplifies the notation considerably, but is not necessary for the method to work [8]. At each mobile the data is queued, until transmitted, in an infinite buffer depending on the source. We are concerned with the allocation of power and/or time and bandwidth to the queues in a queue and channel-state dependent way to assure stability. The capacities of the connecting channels are a correlated random process. Time is divided into small scheduling intervals. At the beginning of the intervals, the capacities (or surrogates such as the S/N ratios) are estimated via pilot signals and this information is then used to make the scheduling decisions during that interval. Using such information can improve the performance dramatically [1]. Owing to the random nature of the arrival and channel processes, the computation or even the existence of stabilizing policies is not at all obvious. The approach is a network extension of the development for the one-node case in [4]. The references [5], [8] are for networks, and contain many extensions to non-unique routing, acknowledgments of receipt of packets required, multicasting, randomly varying numbers of users, randomly available frequencies or opportunistic frequency allocation, etc., and develop methods for getting the a priori routes. This paper uses a simpler Liapunov function perturbation, based on a simple mixing condition, that has many advantages and allows us to deal with processes not covered by previous work. It is more manageable, and it extends the methods so that heavy tailed input processes can be handled. All the extensions in the references can be carried over to the heavy tailed case.

Owing to the non-Markovian nature of the system state[1] classical stability methods cannot be used without revision, and a perturbed Liapunov function method [4], [7] is adapted to obtain the desired results. With this method, and X denoting the vector of queue values (measured in packets) at all the nodes, one starts with a basic Liapunov function $V(X)$ that works for a "mean flow" system. Then one gets a perturbation $\delta V(n)$ to $V(X)$ so that $V(X(n))+\delta V(n)$ can be used as a Liapunov function for the actual non-Markov physical system and imply the desired stability. The actual decision rule is based on the gradient of $V(X)$ and is readily implemented. The basic result is that, if a certain "mean flow" or fluid approximation process is stable, then so is the physical system under our scheduling rule. This stabilizability of the mean flow approximation can often be readily verified. The condition is "nearly" necessary as well.

The $(n+1)$st scheduling interval is called the nth slot.

This work was partially supported by NSF grant DMS-0506928 and ARO contract W911NF-05-10928

[1]For example, Rayleigh fading is not Markovian].

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE **Scheduling and Control of Mobile Communications Networks with Randomly Time Varying Channels by Stability Methods**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Brown University,Division of Applied Mathematics,182 George Street,Providence,RI,02912**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The time argument n denotes the beginning of the nth slot, and is referred to as "time n." Let $X_{i,k}(n)$ denote the queue size at time n at node k of data coming from source i. If node k is not on the path for source i, then $X_{i,k}(n) \equiv 0$. Define the vectors $X_k(n) = \{X_{i,k}(n), i \leq S\}$ and $X(n) = \{X_k(n), k \leq M\}$, with canonical values X_k and X, resp. For weights $w_{i,k} \geq 0$, and until further notice, the basic Liapunov function will be[2]

$$V(X) = \sum_{i,k} w_{i,k} X_{i,k}^p, \quad p \geq 2. \tag{1.1}$$

(1.1) will be modified slightly for the heavy tailed case.

Stability is supposed to assure a robustness of behavior to small changes in the system. For this reason, as well as because $\{X(n)\}$ is rarely Markovian, it is preferable to use methods that do not require Markovianness. The perturbed Liapunov function method is a powerful tool for such a purpose. Generally, there are many criteria that are of interest to each of the users, e.g., mean delay and variance of delay. One should experiment with the form of the Liapunov function to see what the tradeoffs are between competing criteria, a procedure that yields better results than simply working with a single fixed rule, whatever it is.

There is much work on scheduling under various types of randomness. But, other than [4], [5], [8] little is available for the general network case when the channels are randomly varying. For the one-node case, if the rate of transmission is proportional to power, then [1], [10] gets rules for power allocation whose form is similar to ours when $p = 2$ (called "max weight" rules there), and which are based on stability considerations. Large deviations estimates and a Markovian setup are used. Reference [12] considered the problem of dynamic power allocation. Since the channel-rate and data-arrival processes are i.i.d. sequences, the range of applications is small.

The papers [2], [3] deal with related problems, again essentially for one-node systems. There is a set of parallel processors, and the connectivities between the sources and the processors vary randomly. They prove results concerning the limit of (queue length at t)/t, and show that this limit is zero. This is used to show that the integral of the "rates" of transmission per unit time converges. Such a result does not quite imply stability (in our sense) of the queue length process, since it can grow sublinearly. They allocate a single resource (e.g., bandwidth) and the rate is proportional to the allocation. The work [11], for a one node model, has a Markovian channel-state process, the data input sequence is i.i.d., and a "complete resource pooling" condition is required. The decision rule is the same as ours for a quadratic Liapunov function. The emphasis is on stability in the heavy traffic limit, and showing how the problem simplifies there. Section 2 states the assumptions, Section 3 gives the stability proof for the non-heavy-tailed case, and the changes for the heavy-tailed case are in Section 4.

[2]One could let the powers depend on i, k, or use sums of appropriate convex functions [8]. But the simpler form (1.1) is adequate.

II. Assumptions

Queue (i, k) is the queue for source i at node k. If the path for source i does not use node k, then the queue does not exist. Let k denote a canonical node, and $f(i, k)$ the node that the output of queue (i, k) goes to. I.e., queue (i, k) feeds to queue $(i, f(i, k))$. If node k is the final destination for source i, then terms involving $f(i, k)$ are ignored. Let $b(i, k)$ denote the node that queue (i, k) is fed from. I.e., queue $(i, b(i, k))$ feeds to queue (i, k). If node k is the origin node for source i, then terms involving $b(i, k)$ are ignored. Let $\mathcal{F}_n$ denote the minimal σ-algebra that measures the systems data until time n as well as the channel state in slot n. This channel state is available at time n. Let E_n denote the expectation conditioned on $\mathcal{F}_n$. We say that the packets sent in slot n are sent at time n. Let $d_{i,k}(n)$ denote the number of packets sent from queue (i, k) at time n. It will depend on the current channel state and will be a function of the resources (e.g., power, bandwidth) allocated to that queue. Let $a_{i,k}(n)$ denote the random number of arrivals in slot n from the exterior, if any, from source i at node k. These will be non-zero only for the unique node $k(i)$ at which source i enters the network.

Stability. An appropriate definition of stability is a "uniform mean recurrence time" property [4], [5]. Suppose that there are $0 < q_0 < \infty$ and a real-valued $F(\cdot) \geq 0$ such that: For any n, and $\sigma_1 = \min\{k \geq n : |X(k)| \leq q_0\}$, we have[3]

$$E_n [\sigma_1 - n] \leq F(X(n)) I_{\{|X(n)| \geq q_0\}}. \tag{2.1}$$

Then the system is said to be stable. The definition implies recurrence to some compact set. If $|X(n)|$ reaches a level $q_1 > q_0$, then the conditional expectation of the time required to return to a value q_0 or smaller is bounded by a function of q_1, uniformly in the past history and in n. The right side of (2.1) depends only on $X(n)$, and not on any other data, even though the channel and arrival processes are random and correlated.

The decision rule. The number of packets transmitted from queue (i, k) in slot n is $d_{i,k}(n)$, and this depends on the committed resources. The assignments are subject to constraints. If the constraints are only local, such as bounds on the total nodal power, the $d_{i,k}(n)$ for all i can be determined at node k. If the constraints involve more than one node (e.g., if neighboring nodes cannot use the same carrier frequency),

[3]$\sigma_1 = \infty$, unless otherwise defined.

then the assignments require coordination among the nodes.

As in classical stability-control theory, the idea is to choose the $d_{i,k}(n)$ to minimize $E_n V(X(n+1)) - V(X(n))$ as well as possible. To motivate the actual rule, first evaluate $E_n V(X(n+1)) - V(X(n))$. We have

$$\begin{aligned} & w_{i,k}\left[E_n X^p_{i,k}(n+1) - X^p_{i,k}(n)\right] \\ & = w_{i,k} X^{p-1}_{i,k}(n)\left[-d_{i,k}(n) + E_n a_{i,k}(n) + d_{i,b(i,k)}(n)\right] \\ & \qquad + O(X^{p-2}_{i,k}(n)). \end{aligned}$$

Summing over i and k yields, modulo $O(X^{p-2}_{i,k}(n))$ and the "arrival" terms,

$$\begin{aligned} & -\sum_{i,k}\left[w_{i,k}X^{p-1}_{i,k}(n) - w_{i,f(i,k)}X^{p-1}_{i,f(i,k)}(n)\right] d_{i,k}(n) \\ & = -\sum_{i,k} w_{i,k}X^{p-1}_{i,k}(n)\left[d_{i,k}(n) - d_{i,b(i,k)}(n)\right]. \end{aligned} \tag{2.2}$$

The $O(X^{p-2}_{i,k}(n))$ are nonlinear functions of $d_{i,k}(n)$ and higher conditional moments of the $a_{i,k}(n)$, and would be hard to deal with. Fortunately, as in [4], it is enough to work with the term that is first order in the decisions, those in (2.2).

If the decisions can be made independently at each node, then our decision rule (for each k) is a maximizer in

$$\max_{\{d_{i,k}(n):i\}} \sum_i \left[w_{i,k}X^{p-1}_{i,k}(n) - w_{i,f(i,k)}X^{p-1}_{i,f(i,k)}(n)\right] d_{i,k}(n). \tag{2.3}$$

If there are constraints that involve a set of nodes, then the decisions for them must be made together, and the decision rule is a maximizer in

$$\max_{\{d_{i,k}(n);i,k\}} \sum_{i,k} \left[w_{i,k}X^{p-1}_{i,k}(n) - w_{i,f(i,k)}X^{p-1}_{i,f(i,k)}(n)\right] d_{i,k}(n), \tag{2.4}$$

or, equivalently (by rearranging terms in (2.4)), in

$$\max_{\{d_{i,k}(n);i,k\}} \sum_{i,k} w_{i,k}X^{p-1}_{i,k}(n)\left[d_{i,k}(n) - d_{i,b(i,k)}(n)\right]. \tag{2.5}$$

$X(n)$ is rarely Markovian, so classical stochastic stability theory [6] cannot be used directly. However, the perturbed Liapunov function method [4], [7], [9] will allow us to show that the rules (2.3), (2.4), or (2.5) yield a stable system.

Let $L_k(n)$ denote the (vector) set of channel states, at time n, of all of the channels originating at node k, which are the links $\{(i,k),(i,f(i,k)) :$ all i using node $k\}$. $L_k(n)$ could be simply the set of S/N ratios at the receiver corresponding to unit transmitted power. It is convenient to work with the vector $L_k(n)$, rather than with the individual links, since the decisions at each node k depend on the states of all of the outgoing links. $L_k(n)$ might denote other quantities besides the channel quality. For example, there might be power constraints that vary randomly due to interference from exogenous sources. If some link at node k is unavailable at time n, then that fact could also be included in $L_k(n)$. We suppose that the range of values of the channel state vector is a finite set for each node k. We use the (vector-valued) symbol j for the canonical value of $L_k(n)$, for any k,n. The range of the variable j will depend on the node k and will not be specified. Let $u_{i,k}(j,X)$ denote the control function at queue (i,k). It represents the resources (power, time, bandwidth, etc.) allocated to queue (i,k). Also, unless otherwise noted, its dependence on the queues is only on X_k and the required queue values at the immediate upstream nodes, namely the $X_{i,f(i,k)}$ for all i. If source i does not use node k, then ignore $u_{i,k}(j,X)$. The control $u_{i,k}(j,X)$ determines the amount of data that is sent. Let $g_{i,k}(j,X_{i,k},u_{i,k}(j,X))$ denote the actual amount of data that is sent from queue (i,k) under channel state j and control $u_{i,k}(\cdot)$. This defines $d_{i,k}(n)$; i.e., the channel rate for queue (i,k) associated with current channel state $j = L_k(n)$ and control $u_{i,k}(j,X(n))$ is $d_{i,k}(n) = g_{i,k}(j,X_{i,k}(n),u_{i,k}(j,X(n)))$. The $X_{i,k}$ is an argument of $g_{i,k}(\cdot)$ only because the amount sent cannot be larger than the queue content.

Assumptions. (A2.4) requires that there are controls under which the mean service rate/slot at queue (i,k) for any i that uses node k is slightly greater than $\bar{\lambda}^a_i$. Similar conditions are commonly used in the study of the stability of stochastic networks.

A2.1. The maximizing constrained $d_{i,k}(n)$ exist and are Borel functions of the $\{X(n),L_k(n),i,k\}$.

A2.2. There is a $K_1 < \infty$ such that for all i, $E_n|a_{i,k}(n)|^p \le K_1$. There are $\bar{\lambda}^a_{i,k}$ and a $\rho(k)$ which goes to zero as $k \to \infty$ such that $|E_n\alpha_{i,k}(l) - \bar{\lambda}^a_{i,k}| \le \rho(l-n)$, for all $n, l \ge n, \omega, i, k$.

The $\bar{\lambda}^a_{i,k}$, called the mean external data arrival rate for source i at node k, is zero if $k \ne k(i)$. Define $\bar{\lambda}^a_i = \bar{\lambda}^a_{i,k(i)}$.

A2.3. There are $\Pi_{k,j} \ge 0$ such that $\sum_j \Pi_{k,j} = 1$ and $|E_n I_{\{L_k(l)=j\}} - \Pi_{k,j}| \le \rho(l-n)$, for all $n, l \ge n, \omega$.

A2.4. Define $K_0 = \max_{i,k,j,u,X} g_{i,k}(j,X_{i,k},u_{i,k}(j,X))$. There is a control $\{\tilde{u}_{i,k}(\cdot);i,k\}$ under which the following holds. There are $\{\tilde{q}^j_{i,k};i,k\}$ such that $\tilde{q}^j_{i,k} = g_{i,k}(j,X_{i,k}(n),\tilde{u}_{i,k}(j,X(n)))$ if $X_{i,k}(n) \ge K_0$. [4] Also, $g_{i,k}(j,X_{i,k}(n),\tilde{u}_{i,k}(j,X(n))) \le \tilde{q}^j_{i,k}$ if $X_{i,k}(n) < K_0$.

[4]The lower bound K_0 is introduced only because if the queue content is smaller than the maximum of what can be transmitted on a scheduling interval, then the mean output might be too small to assure the $-c_0$ value. E.g., if a queue is empty, then there are no departures.

The $\tilde{q}^j_{i,k}$ satisfy, for nodes k used by source i,

$$\bar{q}_{i,k} = \sum_j \tilde{q}^j_{i,k}\Pi_{k,j} > \bar{\lambda}^a_i. \tag{2.6}$$

Comments on the assumptions. (A2.2) and (A2.3) are simply mixing conditions on the data arrival and channel processes, resp., and do not appear to be restrictive. Let $\Pi_{k,j}$ denote the steady state probability of channel state j at node k. Then (A2.3) says that the conditional probability of state j at time l given the data to time n converges to the steady state value as $l - n \to \infty$. It holds for the received signal power associated with Rayleigh fading.

By the definition (2.6), for $k \neq k(i)$, $\bar{q}_{i,b(i,k)} = \sum_j \tilde{q}^j_{i,b(i,k)}\Pi_{b(i,k),j}$. It is implied by (A2.4) that there is $c_0 > 0$ such that the $\tilde{q}^j_{i,k}$ can be chosen to satisfy

$$\bar{\lambda}^a_i - \bar{q}_{i,k(i)} \leq -c_0, \tag{2.7a}$$

and, for $k \neq k(i)$,

$$\begin{aligned} &\text{average into } (i,k) - \text{average out of } (i,k) \\ &\quad = \bar{q}_{i,b(i,k)} - \bar{q}_{i,k} \leq -c_0. \end{aligned} \tag{2.7b}$$

.

Consider an example, where the control is over either power, bandwidth, or time. Let the rates be proportional to the allocations $B^j_{i,k}$, with constants of proportionality $c^j_{i,k}$. The rate is $q^j_{i,k} = c^j_{i,k}B^j_{i,k}$. There are the resource constraints $\sum_i B^j_{i,k} \leq B_k$ for each j, k, and the mean throughput constraints $\sum_j q^j_{i,k}\Pi_{k,j} > \bar{\lambda}^a_i$, all k. Any solution $q^j_{i,k}$ satisfies (A2.4).

III. Liapunov Function Perturbations and Proof

The perturbations. We now define the Liapunov function perturbation $\delta V(n)$. This will be a sum of terms, one corresponding to each input process one to each output process of each queue. The motivation for the structure of the perturbations should be apparent from the way that they are used in the proof. Additional background, applications and motivation are in [7], [9]. Recall that $k(i)$ = arrival node for data from source i. The "arrival" perturbations are, for $k = k(i)$ and some integer $m \geq 0$,

$$\delta V^a_{i,k}(n) = w_{i,k}X^{p-1}_{i,k}(n)\sum_{l=n}^{n+m-1} E_n\left[a_{i,k}(l) - \bar{\lambda}^a_{i,k}\right], \tag{3.1}$$

The value of m will be chosen in the proof. Set $\delta V^a_{i,k}(n) = 0$ if $k \neq k(i)$.

Recall that the vector-valued channel state j denotes the canonical state of the set of channels on the forward links from the node in question. In (3.2), we define two sets of perturbations. The top one is concerned with the effects of the departure of packets from queue (i,k) on the value of $E_nX^p_{i,k}(n+1) - X^p_{i,k}(n)$, under channel state j, and the "reference" rates $\tilde{q}^j_{i,k}$ of (A2.4). The bottom one is concerned with the effects on this value of the inputs to (i,k) from queue $(i, b(i,k))$, when the channel state at node $b(i,k)$ is j, and under the "reference" rates $\tilde{q}^j_{i,b(i,k)}$. Define

$$\begin{aligned} &\delta V^{d,+}_{i,k,j}(n) = \\ &\quad -w_{i,k}X^{p-1}_{i,k}(n)\tilde{q}^j_{i,k}\sum_{l=n}^{n+m-1} E_n\left[I_{\{L_k(l)=j\}} - \Pi_{k,j}\right], \\ &\delta V^{d,-}_{i,k,j}(n) = \\ &\quad w_{i,k}X^{p-1}_{i,k}(n)\tilde{q}^j_{i,b(i,k)}\sum_{l=n}^{n+m-1} E_n\left[I_{\{L_{b(i,k)}(l)=j\}} - \Pi_{b(i,k),j}\right]. \end{aligned} \tag{3.2}$$

The full Liapunov function perturbation $\delta V(n)$ and the time-dependent Liapunov function $\tilde{V}(n)$ are

$$\begin{aligned} \delta V(n) &= \sum_{i,k}\delta V^a_{i,k}(n) + \sum_{i,k,j,\pm}\delta V^{d,\pm}_{i,k,j}(n), \\ \tilde{V}(n) &= V(X(n)) + \delta V(n). \end{aligned} \tag{3.3}$$

Theorem 3.1. *Under* (A2.1)–(A2.4) *the system is stable.*

Proof. $\tilde{V}(n)$ is the (time-varying) Liapunov function that is to be used. We need to show that there is $c > 0$ such that $E_n\tilde{V}(n+1) - \tilde{V}(n) \leq -c$ when $|X(n)|$ is large enough, and then that this inequality together with the bounds on $\delta V(n)$ imply (2.1).

As usual in stability proofs, one must evaluate

$$\begin{aligned} E_n\tilde{V}(n+1) - \tilde{V}(n) &= \sum_{i,k} w_{i,k}E_n\left[X^p_{i,k}(n+1) - X^p_{i,k}(n)\right] \\ &+ \sum_{i,k} E_n\left[\delta V^a_{i,k}(n+1) - \delta V^a_{i,k}(n)\right] \\ &+ \sum_{i,k,j,\pm} E_n\left[\delta V^{d,\pm}_{i,k,j}(n+1) - \delta V^{d,\pm}_{i,k,j}(n)\right]. \end{aligned}$$

This will be done component by component, and then the results added. This will have the effect of either "averaging" undesirable terms or else suitably dominating them. This is the key to the effectiveness of the method. A first order Taylor expansion yields

$$\begin{aligned} &\sum_{i,k} w_{i,k}E_n\left[X^p_{i,k}(n+1) - X^p_{i,k}(n)\right] = O(X^{p-2}_{i,k}(n)) \\ &+\sum_{i,k} w_{i,k}X^{p-1}_{i,k}(n)\left[E_na_{i,k}(n) - d_{i,k}(n) + d_{i,b(i,k)}(n)\right]. \end{aligned} \tag{3.4}$$

Now consider the "arrival" perturbation component (3.1). If $k = k(i)$, then

$$\begin{aligned} &E_n\delta V^a_{i,k}(n+1) - \delta V^a_{i,k}(n) = \\ &\quad -w_{i,k}X^{p-1}_{i,k}(n)\left[E_na_{i,k}(n) - \bar{\lambda}^a_{i,k}\right] + \epsilon^a_{i,k}(n), \end{aligned}$$

where

$$\epsilon^a_{i,k}(n) = w_{i,k} X^{p-1}_{i,k}(n) \left| E_n a_{i,k}(n+m) - \bar{\lambda}^a_{i,k} \right| \\ +m\, w_{i,k} O(X^{p-2}_{i,k}(n)) \\ \leq w_{i,k} X^{p-1}_{i,k}(n) \rho(m) + m\, O(X^{p-2}_{i,k}(n)). \tag{3.5}$$

Thus, summing over i, k,

$$\sum_{i,k} E_n \left[\delta V^a_{i,k}(n+1) - \delta V^a_{i,k}(n) \right] \\ = -\sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \left[E_n a_{i,k}(n) - \bar{\lambda}^a_{i,k} \right] \tag{3.6}$$

plus error terms bounded by

$$\sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \rho(m) + m \sum_{i,k} O(X^{p-2}_{i,k}(n)). \tag{3.7}$$

By adding (3.4) and (3.6), the $w_{i,k} X^{p-1}_{i,k}(n) E_n a_{i,k}(n)$ terms are replaced by the mean value term $w_{i,k} X^{p-1}_{i,k}(n) \bar{\lambda}^a_{i,k}$ term and an "error" term. The error term will be dominated by the main terms of order $p-1$ for large values of the queue state and appropriate values of m. Such replacements are the motivation for the form of the perturbation (3.1).

Now deal with the top term in (3.2). This will help to "average" the $d_{i,k}(n)$ term in (3.4). By the definitions,

$$E_n \left[\delta V^{d,+}_{i,k,j}(n+1) - \delta V^{d,+}_{i,k,j}(n) \right] = \\ -w_{i,k} E_n X^{p-1}_{i,k}(n+1) \tilde{q}^j_{i,k} \sum_{l=n+1}^{n+m} E_{n+1} \left[I_{\{L_k(l)=j\}} - \Pi_{k,j} \right] \\ +w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,k} \sum_{l=n}^{n+m-1} E_n \left[I_{\{L_k(l)=j\}} - \Pi_{k,j} \right] \tag{3.8}$$

This expression can be written as

$$w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,k} \left[I_{\{L_k(n)=j\}} - \Pi_{k,j} \right] \\ -w_{i,k} E_n X^{p-1}_{i,k}(n+1) \tilde{q}^j_{i,k} \sum_{l=n+1}^{n+m} E_{n+1} \left[I_{\{L_k(l)=j\}} - \Pi_{k,j} \right] \\ +w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,k} \sum_{l=n+1}^{n+m-1} E_n \left[I_{\{L_k(l)=j\}} - \Pi_{k,j} \right]. \tag{3.9}$$

Writing $X^{p-1}_{i,k}(n+1) = X^{p-1}_{i,k}(n) + [X^{p-1}_{i,k}(n+1) - X^{p-1}_{i,k}(n)]$ and expanding the bracketed term yields

$$E_n \left[\delta V^{d,+}_{i,k,j}(n+1) - \delta V^{d,+}_{i,k,j}(n) \right] \\ = w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,k} \left[I_{\{L_k(n)=j\}} - \Pi_{k,j} \right] \tag{3.10}$$

plus error terms that are bounded by

$$w_{i,k} \tilde{q}^j_{i,k} X^{p-1}_{i,k}(n) \left| E_n I_{\{L_k(n+m)=j\}} - \Pi_{k,j} \right| + m\, O(X^{p-2}_{i,k}(n)).$$

Analogously, one can show that

$$E_n \left[\delta V^{d,-}_{i,k,j}(n+1) - \delta V^{d,-}_{i,k,j}(n) \right] = \\ -w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,b(i,k)} \left[I_{\{L_{b(i,k)}(n)=j\}} - \Pi_{b(i,k),j} \right] \tag{3.11}$$

plus error terms bounded by

$$w_{i,k} \tilde{q}^j_{i,b(i,k)} X^{p-1}_{i,k}(n) \left| E_n I_{\{L_{b(i,k)}(n+m)=j\}} - \Pi_{b(i,k),j} \right| \\ +m\, O(X^{p-2}_{i,k}(n)).$$

Adding all terms in (3.4), (3.6), (3.10), (3.11), plus the error terms, and cancelling where possible yields

$$E_n \tilde{V}(n+1) - \tilde{V}(n) = \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \bar{\lambda}^a_{i,k} \\ + \sum_{i,k} \left[-w_{i,k} X^{p-1}_{i,k}(n) d_{i,k}(n) + w_{i,k} X^{p-1}_{i,k}(n) d_{i,b(i,k)}(n) \right] \\ + \sum_{i,k,j} w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,k} \left[I_{\{L_k(n)=j\}} - \Pi_{k,j} \right] \\ - \sum_{i,k,j} w_{i,k} X^{p-1}_{i,k}(n) \tilde{q}^j_{i,b(i,k)} \left[I_{\{L_{b(i,k)}(n)=j\}} - \Pi_{b(i,k),j} \right] \\ + \text{error terms bounded by } (2+2K_0) \times (3.7). \tag{3.12}$$

Separate out the terms in the middle three lines of (3.12) that do not involve the $\Pi_{k,j}$ variables, getting

$$- \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \left[d_{i,k}(n) - d_{i,b(i,k)}(n) \right] \\ + \left\{ \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \sum_j \left[\tilde{q}^j_{i,k} I_{\{L_k(n)=j\}} \right] \right. \\ \left. - \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \sum_j \left[\tilde{q}^j_{i,b(i,k)} I_{\{L_{b(i,k)}(n)=j\}} \right] \right\}.$$

The last expression can be written as

$$- \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \left[d_{i,k}(n) - d_{i,b(i,k)}(n) \right] \\ + \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \left[\tilde{q}^{L_k(n)}_{i,k} - \tilde{q}^{L_{b(i,k)}(n)}_{i,b(i,k)} \right]. \tag{3.13}$$

Suppose, for the moment, that all $X_{i,k}(n) \geq K_0$. Then, by (A2.4) there is a control $\{\tilde{u}_{i,k}(\cdot)\}$ such that under channel state j the output from queue (i,k) will be $\tilde{q}^j_{i,k} = g_{i,k}(j, X_{i,k}(n), \tilde{u}_{i,k}(j, X(n)))$. Since the $d_{i,k}(n)$ are chosen by one of the maximization rules (2.3), (2.4), or (2.5), and the $\tilde{q}^j_{i,k}$ outputs defined in (A2.4) are not necessarily maximizers, the expression (3.13) is non-positive. Using this fact in (3.12) together with the definition of $\bar{q}_{i,k}$ in (A2.4) yields the upper bound to (3.12):

$$\sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \left[\bar{\lambda}^a_{i,k} - \bar{q}_{i,k} + \bar{q}_{i,b(i,k)} \right] \\ +(2+2K_0) \sum_{i,k} w_{i,k} X^{p-1}_{i,k}(n) \rho(m) + m \sum_{i,k} O(X^{p-2}_{i,k}(n)). \tag{3.14}$$

By (2.6), $\bar{\lambda}^a_{i,k} - \bar{q}_{i,k} + \bar{q}_{i,b(i,k)} \leq -c_0$. Select the integer m so that $(2+2K_0)\rho(m) \leq c_0/2$. Thus, since m is now fixed,

$$E_n\tilde{V}(n+1) - \tilde{V}(n) \\ \leq -[c_0/2]\sum_{i,k} w_{i,k}X_{i,k}^{p-1}(n) + O(|X(n)|^{p-2}). \tag{3.15}$$

We also have

$$|\delta V(n)| = O(|X(n)|^{p-1}) \tag{3.16}$$

and, by (3.15),

$$E_n\tilde{V}(n+1) - \tilde{V}(n) \to -\infty, \text{ unif. in } n \text{ as } X(n) \to \infty. \tag{3.17}$$

By (3.17), there are $c_1 > 0$ and $q_0 > 0$, such that, for $|X(n)| \geq q_0$,

$$E_n\tilde{V}(n+1) - \tilde{V}(n) \leq -c_1. \tag{3.18}$$

Given small $\delta > 0$, (3.16) implies that for q_0 sufficiently large,

$$|V(X(n)) - \tilde{V}(n)| \leq \delta(1 + V(X(n)). \tag{3.19}$$

Let σ_0 be a stopping time for which $|X(\sigma_0)| = c_2 > q_0$, and define $\sigma_1 = \min\{n > \sigma_0 : |X(n)| \leq q_0\}$. Then, by (3.18), we have

$$E_{\sigma_0}\tilde{V}(\sigma_1) - \tilde{V}(\sigma_0) \leq -c_1 E_{\sigma_0}[\sigma_1 - \sigma_0]. \tag{3.20}$$

Using (3.20) and the bound (3.19) on $\tilde{V}(n) - V(X(n))$ to bound $\tilde{V}(\sigma_i) - V(X(\sigma_i))$, $i = 0, 1$, yields

$$-\delta E_{\sigma_0}[1 + V(X(\sigma_1))] + E_{\sigma_0}V(X(\sigma_1)) \leq E_{\sigma_0}\tilde{V}(\sigma_1) \\ \leq -c_1 E_{\sigma_0}(\sigma_1 - \sigma_0) + [\delta + V(X(\sigma_0))(1+\delta)]$$

or

$$E_{\sigma_0}(\sigma_1 - \sigma_0) \leq \frac{2\delta + V(X(\sigma_0))(1+\delta) + \delta E_{\sigma_0}V(X(\sigma_1))}{c_1},$$

which implies that the definition of stability (2.1) holds since $V(X(\sigma_1)) \leq \sup_{|x|\leq q_0} V(x)$.

There is no space to complete the details when some components of $X(n)$ are less than K_0. The required adjustments are minor and it can be shown, similarly to what was done in [8], that the contributions of those terms are bounded, so that the contributions for large $X_{i,k}(n)$ dominate. The details are omitted. ■

Comments. The rule (2.3) requires that each node k know the value of the $X_{i,k}(n)$ and $X_{i,f(i,k)}(n)$ for all i that use node k. It is easily seen from the proof that the value of $X_{i,f(i,k)}(n)$ need only be known approximately at node k. The quantities q_0, δ, c_1 all depend on the values of the $\rho(\cdot)$ in (A2.2) and (A2.3) and on the excess capacity of the system as quantified by c_0. If the rate of mixing of the channel process is very slow, then the queues will often have very large excursions, despite the fact of stability. Consider the possibility that some links are preempted by priority users from time to time, where the intervals of availability are defined by a renewal process that is independent of the arrival and channel rate processes. Then it can be shown that the results continue to hold, but with the $\bar{q}_{i,k}$ multiplied by the fraction of time that the channel is available, so the capacity must be sufficient to handle the down times. Under the other assumptions, (A2.4) is sufficient but not necessary for stability, but it is "nearly" necessary in the following sense. Suppose that for each allowable choice of the $\{\tilde{q}_{i,k}^j\}$, there is some (i_0, k_0) such that $\bar{q}_{i_0,b(i_0,k_0)} - \bar{q}_{i_0,k_0} > 0$. Then the system is not stable.

IV. Heavy Tailed Input Distributions

Poisson arrival processes. Suppose that source i generates input files whose lengths have heavy tailed distributions in that their second moment is finite but for some $0 < \delta < 1$ the $(1+\delta)$th moment is finite. Let the file creation process be Poisson with rate $\bar{c}_i$/slot with mean file size $\bar{m}_i$. Thus the mean input rate/slot is $\bar{d}_i = \bar{c}_i\bar{m}_i$. Suppose that successive files are mutually independent, and that the file size is known when created[5] and that the file is sent to the system at a rate $\bar{\lambda}_i^a > \bar{d}_i$ packets/slot. [Note the new definition of $\bar{\lambda}_i^a$.] Use the following slightly revised model of the system. Introduce a fictitious node, called $p(i)$. Suppose that the entire contents of any new file from source i is sent all at once to this node, and then sent from this node to the true entry node $k(i)$ of the network at a rate $\bar{\lambda}_i^a$, if queue $p(i)$ is not empty. This is the source i process that would have gone to node $k(i)$ directly if node $p(i)$ were not introduced, as in the previous sections.

Augment $X(n)$ by the new state $X_{p(i)}(n)$ and add the term $[1 + X_{p(i)}]^{1+\delta}/(1+\delta)$ to $V(X)$. This term will not affect the decision rule since the output from node $p(i)$ is not controllable. From the point of view of the proof, the input rate to node $k(i)$ from $p(i)$ is constant at $\bar{\lambda}_i^a$ if $X_{p(i)}(n) > 0$. Suppose, w.l.o.g., that the input file size is always greater than $\bar{\lambda}_i^a$ and is an integral multiple of $\bar{\lambda}_i^a$.

The only new contribution to the proof is the behavior of (divided by $1+\delta$)

$$E_n[1 + X_{p(i)}(n+1)]^{1+\delta} - E_n[1 + X_{p(i)}(n)]^{1+\delta}. \tag{4.1}$$

This can be written as

$$E_n\left[1 + X_{p(i)}(n) + \theta_i(n)\left(a_{p(i)}(n) - d_{p(i)}(n)\right)\right]^\delta \\ \times \left(a_{p(i)}(n) - d_{p(i)}(n)\right), \tag{4.2}$$

where $\theta_i(n)$ is a random variable with values in $[0,1]$, and $d_{p(i)}(n) = \bar{\lambda}_i^a$ if the queue is not empty. In the event that there is no arrival at n, by the monotonicity

[5]These file sizes will not be used by the decision rule. It is only required that they be included in $\mathcal{F}_n$,

of $(1+x)^\delta$ we can bound (4.2) from above by

$$-d_{p(i)}(n)\left[1+X_{p(i)}(n)\right]^\delta. \tag{4.3}$$

In the event that there is an arrival at n, treating the two components due to the arrivals and departures separately and using $(a+b)^\delta \le a^\delta + b^\delta$, (4.2) can be bounded above by

$$\begin{aligned} &-d_{p(i)}(n)\left[1+X_{p(i)}(n)\right]^\delta \\ &\quad + \left[1+X_{p(i)}(n)\right]^\delta \bar{m}_i + E|a_{p(i)}(n)|^{1+\delta}. \end{aligned} \tag{4.4}$$

To get the desired bound, weigh (4.3) by $(1-\bar{c}_i)$ and (4.4) by $\bar{c}_i$. Repeat the above procedure for each heavy tailed input. The rest of the proof is as for Theorem 3.1. One only needs to keep track of the different powers of the $X_{i,k}(n)$. The condition (A2.2) need hold only for the non-heavy=tailed inputs. No additional perturbations are needed, due to the Poisson and i.i.d property of the input sequence $a_{p(i)}(n)$.

Renewal arrival processes. Continue to suppose that the heavy tailed arrival file sizes from source i are i.i.d., with mean $\bar{m}_i$. But now, let the arrival times constitute a random point process. Let $I_i^a(l)$ denote the indicator function of an arrival at time l, and suppose that, for some $\bar{c}_i > 0$, $|E_n I_i^a(l) - \bar{c}_i| \le \rho(l-n)$ for all $n, l \ge n, \omega$, where $\rho(n) \to 0$ as $n \to \infty$. Then the proof, modified as above, goes through, with the additional perturbation

$$\delta V_{p(i)}^a(n) = \left[1+X_{p(i)}(n)\right]^\delta \bar{m}_i \sum_{l=n}^{n+m-1} \left[E_n I_i^a(l) - \bar{c}_i\right], \tag{4.5}$$

where the integer m is chosen as in the proof. Next, suppose that the file sizes are correlated and that $|E_n I_i^a(l) a_{p(i)}(l) - \bar{d}_i| \le \rho(l-n)$ for all $n, l \ge n, \omega$. then the modified proof goes through, with use of the perturbation

$$\begin{aligned} &\delta V_{p(i)}^a(n) \\ &= \left[1+X_{p(i)}(n)\right]^\delta \sum_{l=n}^{n+m-1} \left[E_n I_i^a(l) a_{p(i)}(l) - \bar{d}_i\right]. \end{aligned} \tag{4.6}$$

.

References

[1] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting. Providing quality of service over a shared wireless link. *IEEE Communications Magazine*, 2001.

[2] N. Bambos and G. Michailidis. Queueing and scheduling in random environments. *Adv. in Appl. Prob.*, 36:293–317, 2004.

[3] N. Bambos and G. Michailidis. Queueing dynamics of random link topology: Stationary dynamics of maximal throughput schedules. *Queueing Systems*, 50:5–52, 2004.

[4] R. Buche and H.J.Kushner. Control of mobile communication systems with time-varying channels via stability methods. *IEEE Trans on Autom. Contr.*, 49:1954–1962, 2004.

[5] H. J. Kushner. Control of multi-node mobile communications networks with time varying channels via stability methods. submitted.

[6] H.J. Kushner. *Stochastic Stability and Control.* Academic Press, New York, 1967.

[7] H.J. Kushner. *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory.* MIT Press, Cambridge, Mass., 1984.

[8] H.J. Kushner. Scheduling and control of multi-node mobile communications systems with randomly-varying channels by stability methods. To appear in the proceedings of an IMA workshop on wireless communications, June 2005, to be published by Springer, 2006.

[9] H.J. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications.* Springer-Verlag, Berlin and New York, 2003.

[10] S. Shakkoti and A. Stolyar. Scheduling for multiple flows sharing a time-varying channel: The exponential rule. In M Suhov, editor, *Analytic Methods in Applied Probability: In Memory of Fridrih Karpelevich, American Math. Soc. Transl. , Series 2, Volume 207*, pages 185–202. American Mathematical Society, Providence, 2002.

[11] S. Stolyar. Max weight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. *Ann. of Appl. Probab*, 14:1–53, 2004.

[12] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Trans. Automatic Control*, 39:466–478, 1993.