Fluid Models of Wireless Networks

Atilla Eryilmaz
Department of Electrical and Computer Engineering
and
Coordinated Science Lab
University of Illinois at Urbana-Champaign
1308 W. Main Street
Urbana, IL 61801
Email: eryilmaz@uiuc.edu

R. Srikant
Department of General Engineering
and
Coordinated Science Lab
University of Illinois at Urbana-Champaign
1308 W. Main Street
Urbana, IL 61801
Email: rsrikant@uiuc.edu

***Abstract*— Given a wireless network that uses a certain, possibly complex strategy to manage the transmission of flows among its users, our goal is to understand its behavior by studying a *fluid approximation*. The validity and ease of working with a fluid model, instead of the original system is demonstrated by reviewing such models for different allocation strategies in both the single-hop and multi-hop setting. Simulation results in a simple case show that the approximate model successfully captures the behavior of the original system reasonably well.**

I. Introduction

In a wireless network, the information transmission rate (which we often refer to as the *service rate*) over each link is time-varying due to fading and interference from other links. In order to increase capacity, users participate in forwarding other users' data over the resources available to them.This allows for transmission with low power consumption and interference. On the other hand, such a cooperation requires a set of rules that control the access of the users to the transmission medium. Such medium access strategies are called *scheduling* policies.

A *scheduling policy* is an allocation of service rates to the various flows, under the constraint that, at each time instant and each channel state, the set of allocated rates lies within some allowable set of rates. In this paper, we assume that the rate of arrivals of the flows to the different users, the variation of the wireless channel characteristics as well as the service rates allocated to flows under a scheduling policy are all well-described by stochastic processes.

In this work, we aim to study the throughput of scheduling strategies, which stabilize the system using only queue length information and the current channel state(i.e. without knowing the channel or arrival statistics). A number of such policies have been provided in the networking literature [6],[7], [11], [10], [9].

In general, understanding the exact behaviour of the stochastic system is tedious and unnecessary. Instead, we want to examine an equivalent system that is much easier to analyze and simulate, and that provides a rough estimate on how the original system would operate. This leads to the *fluid approximation* of the queues.

For any queueing system, the number of packets in the queue is a stochastic process with discontinuous jumps. We observe that when the queue sizes are large, it is reasonable to replace these discontinuities by smooth continuous functions of time. Fluid models are among the simplest models to work with, due to the fact that deterministic methods, such as Lyapunov techniques, can be applied to study their behavior. The validity of using fluid model approximations for stability analysis of stochastic networks has been studied by Dai [4], Chen [3], Meyn [5], Hajek [8] under various assumptions on the arrival and the service processes.

The rest of the paper is organized as follows. In Section II, we start by considering a multi-user, single-hop wireless network (which frequently appears in the uplink or downlink cellular model), and provide an analysis of the equivalent fluid model under a particular scheduling policy. Section III generalizes the model to a multi-hop scenario and reviews the scheduling policy introduced in [9]. In Section IV, we present results comparing fluid and stochastic simulations of a simple wireless network (with one transmitter and two receivers). Finally, Section V provides the concluding remarks and future directions.

II. Single-hop Transmission Model

Consider a wireless network where N transmissions take place over a single time-varying channel. Examples of such a network can be a single transmitter simultaneously sending data to N receivers (as in the downlink of a cellular system), or N transmitters sending data to a single receiver (as in the uplink of a cellular system). Both of these cases are depicted in Figures 1(a), 1(b).

We assume that the arriving bits are stored in N separate queues, one for each data stream. Assuming that time is slotted, the evolution of the size i^{th} queue is governed by the following equation:

$$x_i[k+1] = (x_i[k] + a_i[k] - \eta_i[k])^+ ,$$

where $a_i[k]$ is the number of bits arriving to Queue-i during slot k and $\eta_i[k]$ is the number of bits from Queue-i that are served within slot k. Also, $(y)^+$ equals y when $y \geq 0$ and equals 0 otherwise.

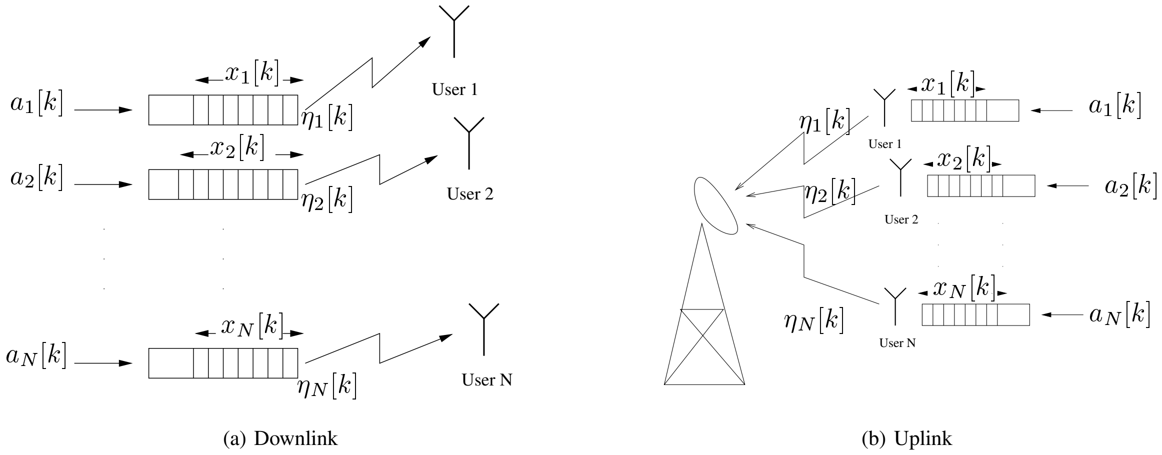


Fig. 1. Examples of wireless networks that fit into our description

The state of the wireless channel is assumed to be fixed within a time slot, but is allowed to vary from one slot to the next. We let J denote the number of possible channel states. Suppose that the channel is in State j at time k, then $\{\eta_i[k]\}$ is constrained to be in the corresponding N dimensional region S_j. We assume that S_j is a bounded, convex region. We note that, the convexity assumption can easily be relaxed by taking the convex hull of each region $\{S_j\}$ and working with them.

A. Scheduling Policy

The class of scheduling policies we wish to study in this section is described as follows. At any time k, given the current channel state $s[k]$, the scheduler chooses a service rate vector $\vec{\mu} = (\mu_1, \cdots, \mu_N)' \in S_{s[k]}$ that satisfies

$$\vec{\mu}[k] \in \arg \max_{\vec{\eta} \in S_{s[k]}} \sum_{i=1}^{N} x_i[k]\eta_i \tag{1}$$

As an example, the rate allocation according to the above strategy for the two user broadcast channel scenario is provided in Figure 2.

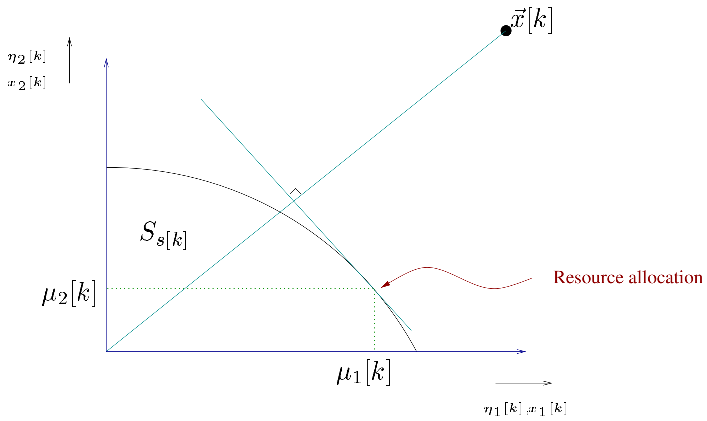


Fig. 2. An implementation of the policy

B. Fluid Model Approximation

In this model, time is no longer assumed to be discrete but continuous. Also, we approximate the arrival process to the i^{th} queue by a fluid with a constant arrival rate, where the constant arrival rate is taken to be the mean arrival rate of the stochastic arrival process. Let λ_i denote this arrival rate. Let us define the mean allowable rate region as

$$\bar{S} = \{\vec{\eta} : \vec{\eta} = \sum_{j=1}^{J} \pi_j \vec{\eta}^{(j)},\ \vec{\eta}^{(j)} \in S_j\ \forall j\},$$

where π_j is the stationary probability of the channel state being j.

Then, the evolution of the queue lengths can be described by the following differential equation

$$\dot{x}_i(t) = \begin{cases} \lambda_i - \bar{\mu}_i(t), & x_i(t) > 0 \\ (\lambda_i - \bar{\mu}_i(t))^+, & x_i(t) = 0 \end{cases} \qquad i = 1, \cdots, N, \tag{2}$$

where $\bar{\mu}(t)$ is chosen such that it satisfies

$$\vec{\bar{\mu}}(t) \in \arg \max_{\vec{\eta} \in \bar{S}} \sum_{i=1}^{N} x_i(t)\eta_i. \tag{3}$$

Notice the difference between the original scheduling policy described by (1) and (3). In the approximate version, the fluctuations due to arrivals as well as the channel variations are replaced by their means. Next, we prove the stability of the approximate model using Lyapunov analysis.

1) Stability Analysis: Consider the following Lyapunov function:

$$V(\vec{x}) = \frac{1}{2} \sum_{i=1}^{N} x_i^2.$$

For stability, it is sufficient to show that the time derivative of the Lyapunov function is strictly negative whenever $\|x\| := \sqrt{\sum_{i=1}^{N} x_i^2}$ is larger than some finite (but possibly large) number. Basically, this amounts to showing that the queue lengths will tend to decrease if they are sufficiently large.

Theorem 1: If the vector of arrival rates, $\vec{\lambda} = (\lambda_1, \cdots, \lambda_N)'$ lies inside the mean allowable rate region, $\bar{S}$, then there exists some $\delta > 0$, such that

$$\dot{V}(\vec{x}) \leq -\|x\|\delta,$$

under the policy described by (3).

Proof: Observe that

$$\dot{V}(\vec{x}) = \sum_{i=1}^{N} x_i(\lambda_i - \bar{\mu_i}).$$

We can write $x_i = \|x\| \cos(\theta_i)$ for appropriate values of θ_i and expand the previous equation as follows.

$$\dot{V}(\vec{x}) = \|x\| \sum_{i=1}^{N} (\lambda_i - \bar{\mu}_i) \cos(\theta_i), \tag{4}$$

where, due to (3) and $\|x\| \geq 0$ being a constant for a fixed $\vec{x}$, we have

$$\vec{\bar{\mu}} \in \arg\max_{\vec{\eta} \in \bar{S}} \sum_{i=1}^{N} \cos(\theta_i)\eta_i.$$

It can be shown that the summation of (4) is determined by the distance between two parallel hyperplanes passing through $\vec{\lambda}$ and $\vec{\bar{\mu}}$, where the slope of the hyperplanes is determined by $\vec{x}$. Also , it is true that the summation is upper-bounded by a negative constant $-\delta$. Hence, we obtain

$$\dot{V}(\vec{x}) \leq -\|x\|\delta.$$

■

Clearly, the approximate model analysis is extremely simple compared to stochastic model. In [6], [7], a more general class of scheduling policies is proved to be stabilizing for the stochastic system. The stochastic analysis proves that the mean queue lengths are bounded, but the key intuition can be obtained by simply studying the fluid model as above.

III. Multi-hop Transmission Model

As briefly described in Section I, in the multi-hop transmission model, each user is not only responsible for transmitting its own data, but also for forwarding flows that pass through it.

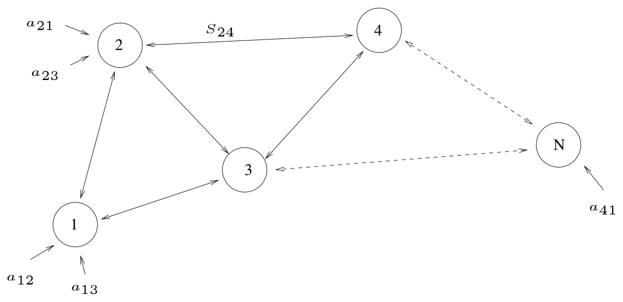


Fig. 3. Multi-hop wireless network model

In this scenario, we model the wireless network as a graph with N nodes and E edges an in Figure 3. Each node corresponds to a user and each edge represents a wireless link. The state of the link between users i and j is denoted by S_{ij}. $\underline{S}$ denotes the matrix of allowable rates as a function of the channel states at the time. The arrival for the flow that originated at node i and is destined for node j is denoted with a_{ij}.

At each node, $N-1$ queues are maintained, one for each destination. The flows that arrive at node i, that is destined for node j is kept in Queue$(i \rightarrow j)$ and $x_i^{(j)}$ is used to denote the size of this queue.

Similar to our definition of the mean allowable rate region in Section II, we define the matrix of mean allowable rate regions as

$$\bar{\underline{S}} = \sum_{\underline{S}} \pi_{\underline{S}}\, \underline{S}.$$

A. Scheduling Policy

For the described model, we will analyze the scheduling policy introduced by Neely, et al. [9]:

1) For link (i,j), at time k, find

$$c^*_{ij}[k] = \arg\max_{c \in \{1,\cdots,N\}} \left(x_i^{(c)}[k] - x_j^{(c)}[k] \right),$$

and let

$$W^*_{ij}[k] = \max\left\{ \left(x_i^{(c^*_{ij})}[k] - x_j^{(c^*_{ij})}[k] \right), 0 \right\}.$$

2) Choose the service rate matrix $\underline{\mu}[k]$ such that

$$\underline{\mu}[k] \in \arg\max_{\underline{\eta} \in \underline{S}} \sum_{i,j} \eta_{ij} W^*_{ij}[k] \tag{5}$$

Next, we will approximate this system with a fluid model and justify its stability with a simple calculation.

B. Fluid Model Approximation

Just like in the previous approximation, we approximate the arrival and channel processes with their means. The queue length evolution of each queue is governed by a differential equation of the form (2). And finally at any time t, the service rates are chosen from the *mean* allowable rate matrix according to (5).

1) Stability Analysis: To study the stability of the fluid model, we begin by defining the following Lyapunov function:

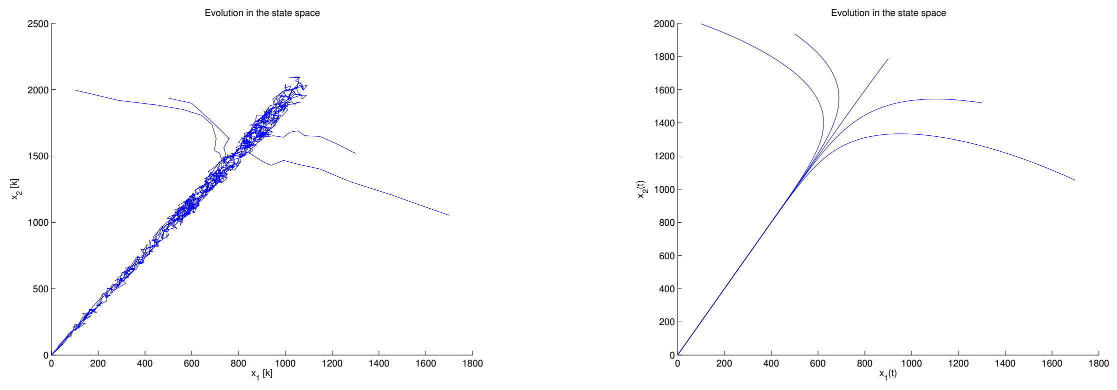
$$W(\underline{x}) = \frac{1}{2} \sum_{i,c} \left(x_i^{(c)} \right)^2.$$

The next step is to examine the derivative of $W(\underline{x}(t))$:

$$\dot{W}(\underline{x}(t)) = \sum_{i,c} \dot{x}_i^{(c)}(t) x_i^{(c)}(t). \tag{6}$$

Notice that $\dot{x}_i^{(c)}(t)$ represents the rate of change in the size of Queue$(i \rightarrow c)$, which can be written as the incoming rate of flows into node i that are destined for node c (including the exogeneous arrivals) minus the outgoing rate of flows also destined for node c, i.e.,

$$\dot{x}_i^{(c)}(t) = \sum_j \mu_{ji}^{(c)}(t) + \lambda_{ic} - \sum_k \mu_{ik}^{(c)}(t)$$



(a) Stochastic model

(b) Fluid model

Fig. 4. Simulations of the models with various initial starting points

Substituting this equation into (6) yields

$$\begin{aligned}\dot{W}(\underline{x}(t)) &= \sum_{i,c}\left(\lambda_{ic}+\sum_{j}\mu_{ji}^{(c)}(t)-\sum_{k}\mu_{ik}^{(c)}(t)\right)x_i^{(c)}(t) \\ &= \sum_{i,c}\lambda_{ic}x_i^{(c)}(t) \\ &\quad -\sum_{j,k}\sum_{c}\mu_{jk}^{(c)}(t)\left(x_j^{(c)}(t)-x_k^{(c)}(t)\right). \qquad (7)\end{aligned}$$

Notice the resemblence between (5) and (7). Equation (5) chooses the service rate matrix that maximizes the weighted sum, where the weight is equal to the maximum value of $\left(x_j^{(c)}(t)-x_k^{(c)}(t)\right)$ over all possible c. Using this observation, it can be shown that the overall expression must be negative provided that the mean arrival rate vectors lie within the corresponding mean allowable rate regions.

As shown in [9], the stability of the stochastic model can be established using law of large number type of assumptions on the arrival and channel processes.

IV. Simulations

Motivated by the stability analysis in the previous sections, in this section, we compare the behavior of a fluid model and a stochastic model of a wireless network through simulations. Consider a downlink transmission scenario with 2 users and a wireless channel having 5 states. The arrivals to both of the queues are Poisson, with their mean lying within but close to the boundary of the average allowable rate region.

Our goal is to compare the behaviors of the original system and its fluid approximation. For this pupose, we simulated the stochastic system and the fluid model, starting from various initial values of the queue lengths, $(x_1(0), x_2(0))$. Figure 4 depicts the results.

It can be seen from the figure that except for fluctuations, due to arrival and channel variations, the queue length evolution of the fluid model successfully mimics the queue length evolution of the stochastic model independent of the initial conditions.

This suggests that behavior of the queues under various scheduling policies may be well approximated by the behavior of the fluid model, when the traffic is close to capacity.

V. Conclusion and Further Directions

In this paper, we reviewed recent literature on the usefulness of fluid approximations in studying wireless networks operating under scheduling policies that exploit queue length and channel state information. We considered scheduling policies for both the single-hop and the multi-hop network scenarios. We then showed that stability can be established by examining appropriate fluid models.

This work motivates the usage of fluid models as a tool for obtaining an approximate estimate on the capacity of a wireless network. A further direction may be to exploit this tool for determining the capacity region of an actual network.

An example of such a network would be the popular IEEE 802.11 wireless network. Simulation of a network that utilizes the 802.11 protocol is impractical, especially when the number of nodes increase. Instead, following the analysis of Cali et al.[1], [2], we can assume that after each collision, the backoff interval is chosen according to a geometric distribution with mean equal to the actual backoff interval. This way, we avoid the complications of the back-off behavior of 802.11, and we assume that each user will attempt to transmit in a slot with a fixed probability, say p. Then, if we further approximate the model as a fluid (like we did in the previous sections), then given a set of mean flow rates, we can simulate this approximate model, in order to get a reasonable estimate on the 802.11 network. The challenge here is to find an appropriate fluid model for this purpose.

Acknowledgment

This work was supported by DARPA Grant F30602-00-2-0542.

References

[1] F. Cali, M. Conti, and E. Gregori. IEEE 802.11 wireless lan: capacity analysis and protocol enhancement. *Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM*, 1:142–149, 1998.

[2] F. Cali, M. Conti, and E. Gregori. IEEE 802.11 protocol: design and performance evaluation of an adaptive backoff mechanism. *IEEE Journal on Selected Areas in Communications*, 18(9):1774–1786, September 2000.

[3] H. Chen. Fluid approximations and stability of multiclass queueing networks: work conserving disciplines. *Annals of Applied Probability*, 5:637–665, 1995.

[4] J. G. Dai. On the positive harris recurrence for multiclass queueing networks: A unified approach via fluid limit models. *Annals of Applied Probability*, pages 49–77, 1995.

[5] J. G. Dai and S. P. Meyn. Convergence and moments of networks and their fluid models. *IEEE Transactions on Automatic Control*, 40:1889–1904, 1995.

[6] A. Eryilmaz, R. Srikant, and J. Perkins. Stable scheduling policies for fading wireless channels. Technical Report. Available at http://www.comm.csl.uiuc.edu/˜srikant.

[7] A. Eryilmaz, R. Srikant, and J.R. Perkins. Stable scheduling policies for fading channels. Yokohama, Japan, July 2003. To be presented at IEEE International Symposium on Information Theory, ISIT.

[8] B. Hajek. Large bursts do not cause instability. *IEEE Transactions on Automatic Control*, 45:116–118, January 2000.

[9] M.J. Neely, E. Modiano, and C.E. Rohrs. Dynamic power allocation and routing for time varying wireless networks. *Proceedings of IEEE Infocom*, April 2003. To appear.

[10] S. Shakkottai and A. Stolyar. Scheduling for multiple flows sharing a time-varying channel: The exponential rule. *Translations of the AMS,* A volume in memory of F. Karpelevich, 2002. To appear.

[11] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, pages 1936–1948, December 1992.

Biographies of the Authors

Atilla Eryilmaz received his B.S. from Bogazici University, Istanbul in 1999, his M.S. from the University of Illinois at Urbana-Champaign in 2001. He is continuing his Ph.D. studies at the University of Illinois as a research assistant in the Coordinated Science Laboratory. His main research interests are the fields of communication networks, information theory and queueing theory.

R. Srikant received his B.Tech. from the Indian Institute of Technology, Madras in 1985, his M.S. and Ph.D. from the University of Illinois in 1988 and 1991, respectively, all in Electrical Engineering. He was a Member of Technical Staff at AT&T Bell Laboratories from 1991 to 1995. He is currently with the University of Illinois, where he is an Associate Professor in the Department of General Engineering, a Research Associate Professor in the Coordinated Science Lab and an Affiliate in the Department of Electrical and Computer Engineering. He was an associate editor of Automatica, and is currently on the editorial boards of the IEEE/ACM Transactions on Networking and IEEE Transactions on Automatic Control. He was the chair of the 2002 IEEE Computer Communications Workshop in Santa Fe, NM. He received an NSF CAREER award in 1997. His research interests include communication networks, stochastic processes, queueing theory, information theory, control theory and game theory.