Robust Rate Control for Heterogeneous Network Access in Multi-Homed Environments

Tansu Alpcan, *Member, IEEE,* Jatinder Pal Singh, *Member, IEEE,* and Tamer Başar, *Fellow, IEEE*

Abstract—We investigate a novel robust flow control framework for heterogeneous network access by devices with multi-homing capabilities. Towards this end, we develop an H^∞-optimal control formulation for allocating rates to devices on multiple access networks with heterogeneous time-varying characteristics. H^∞ analysis and design allow for the coupling between different devices to be relaxed by treating the dynamics for each device as independent of the others. Thus, the distributed end-to-end rate control scheme proposed in this work relies on minimum information and achieves fair and robust rate allocation for the devices. An efficient utilization of the access networks is established through an equilibrium analysis in the static case. We perform measurement tests to collect traces of the available bandwidth on various WLANs and Ethernet. Through simulations, our approach is compared with AIMD and LQG schemes. In addition, the efficiency, fairness, and robustness of the H^∞-optimal rate controller developed are demonstrated via simulations using the measured real world network characteristics. Its favorable characteristics and general nature indicate applicability of this framework to a variety of networked systems for flow control.

Index Terms—Wireless communication, heterogeneous networks, multi-homing, rate control, H^∞-optimal control.

1 Introduction

Contemporary networks are heterogeneous in their attributes such as the supporting infrastructure, protocols, and offered data rates. The multitude and variety of existing and emerging wireless and wired networking technologies continue to be the driving force towards convergence of networks. It is commonplace today to have electronic devices with multiple networking capabilities. Personal computing devices, e.g., laptops, UMPCs (ultra-mobile PCs), PDAs, smartphones, are typically equipped with several access systems ranging from different types of IEEE 802.11 wireless local area networks (WLAN) to Ethernet, GPRS, and UMTS.

On end-user devices a variety of applications emerge with different bandwidth requirements for multimedia access, gaming, and collaboration. Our objective is efficient utilization of multiple networks by devices via rate control and optimal assignment of traffic flows to available networks. The functionality we envision can be described in a hypothetical scenario as follows. Imagine a user in a corporate setting participating in a video conference call via a device having both Ethernet and WLAN (say IEEE 802.11g) connectivity. While engaged in the conference proceedings the user is retrieving relevant files from a remote server, and at the same time generating content by taking notes on her/his blog. Several traffic flows of different characteristics are hence created by the device, which dynamically monitors the networks at its disposal. The device then routes the flows via these networks and dynamically reassigns them to different networks based on the varying network characteristics like available bit rate (ABR) and delay.

While the distribution of traffic flows amongst different networks can enable better network utilization than single network use at a time, the variability in network characteristics like ABR and delay makes the problem of flow control and assignment challenging. This and similar problems have been explored from different perspectives. A game theoretic framework for bandwidth allocation for elastic services in networks with fixed capacities has been addressed in [1]–[3]. Packet scheduling for utilization of multiple networks has been investigated in [4]. A solution for addressing the handoff, network selection, and autonomic computation for integration of heterogeneous wireless networks has been presented in [5], which, however, does not address efficient simultaneous use of heterogeneous networks and does not consider wireline settings. In [6], the authors have explored design of a network comprised of wide area and local area technologies where user devices select among the two technologies in a greedy fashion so as to maximize a utility function based on wireless link quality, network congestion, etc. In [7], an optimal rate control scheme has been investigated where the queue size at a bottleneck link is optimized using an H^∞ control formulation.

As evidenced by the existing body of literature, multiple network access in multi-homed settings where devices simultaneously utilize available networks is a problem of keen academic interest. A few of the recent efforts include the work by Thompson et. al. [8] where flow based multi-homing in residential environments with access to heterogeneous networks has been explored, Shakkottai et. al. [9] where multi-homing of IEEE 802.11 wireless devices to access multiple APs has

- *T. Alpcan is with the Deutsche Telekom Laboratories, Technische Universität Berlin, Germany.*
  E-mail: see http://www.tansu.alpcan.org
- *J. P. Singh is with the Dept. of Electrical Engineering at Stanford University and Deutsche Telekom Laboratories, CA.*
  E-mail: jatinder.singh@telekom.de
- *T. Başar is with the Coordinated Science Laboratory, University of Illinois, Urbana, IL.*
  E-mail: tbasar@control.csl.uiuc.edu

Research supported by Deutsche Telekom AG.
An earlier version of this paper has appeared in the Proceedings of WiOpt 2007, 5th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, April 16-20, Limassol, Cyprus.

been investigated, Kumar et. al. [10] where globally optimal user-network association in an integrated WLAN and UMTS hybrid cell is formulated as a generic Markov decision process connection routing decision problem, and the work by Chandra et. al. [11] where system-level recommendations and designs for such multi-homing in wireless LANs have been addressed. Recently, we have investigated the problem in [12] and [13].

In this paper, we address the problem of optimal rate control and assignment of flows on a device onto multiple networks with randomly varying characteristics. As a first step, we focus on flow control and leave other important issues such as power control to future research. The flow control scheme we develop has several favorable properties including decentralized nature, minimum information requirement, efficiency in terms of bandwidth utilization, fairness, and robustness. Furthermore, it is applicable to a variety of networked systems for flow control. Here, we specifically use it in developing principles for an envisioned *middleware* functionality in the context of heterogeneous network access for multi-homed devices.

The main motivation for our approach stems from the practically observed varying characteristics of networks. We consider a setting where multiple devices hosting a variety of flows with different requirements have access to several different access networks. Each device has to decide on its own flow rate on each network, which it determines by measuring the available bandwidth and taking into account delay characteristics. It is worth noting that during this decision process, each device has access only to its own limited observations and not to the ones of other devices. This distributed rate control scheme should at the same time meet performance objectives such as full utilization of the available bandwidth on the access networks, robustness with respect to ABR variations, measurement noise and delay, and prevention of excessive rate fluctuations leading to instabilities and jitter. Furthermore, trade-offs such as preference for a particular network due to its favorable delay characteristics need to be taken into account.

Unlike our recent work [14], the devices here do not capture the characteristics of the network, which vary randomly over time within a Markov model, but rather use a linear state-space system to keep track of current and past observations on available bandwidth. We consider a worst-case formulation and let the devices update their flow rates under limited-information obtained via measurements. The H^∞-optimal control theory provides a solid quantitative formulation for this purpose with its robustness and general applicability. The use of H^∞ analysis allows us to treat the dynamics for each device as independent of the others and the variations in the available bandwidth are modeled as unknown disturbances. The problem of bandwidth utilization thus becomes one of disturbance rejection which is addressed by H^∞-optimal control using a linear feedback control mechanism. We perform simulations in Matlab and demonstrate that the H-infinity-based rate control leads to high bandwidth utilization in multi-homed scenarios with dynamically varying characteristics of IEEE 802.11b, IEEE 802.11g, and Ethernet networks. Furthermore, we analytically prove the stability of our rate control for fixed capacities and demonstrate, via simulations, its robustness under varying network utilization scenarios.

The rest of the paper is organized as follows: We describe the network access problem, system description, and our measurement results in Section 2. The network model is presented in Section 3. The analytical framework for robust flow control is discussed in Section 4, which also includes an equilibrium analysis and an illustrative example. Section 5 generalizes the formulation in Subsection 4.2. We describe the simulations and results obtained therefrom in Section 6. The paper concludes with remarks and a discussion of future research directions in Section 7.

2 System Description and Measurements

We describe the underlying architecture for routing and rate control of flows originating from applications running on a device via access networks that are available to the device. Essential to the design is a *middleware* (Figure 1) that runs a lightweight tool to estimate the uplink and downlink bandwidths and delay for the flows on different networks. Applications running on the device consult the middleware for routing of flows to and from the corresponding hosts in the Internet via different networks. For the uplink flows, network assignment is done by assigning flows generated by the applications on the device to suitable networks by consulting the middleware. For the downlink flows, suitable network assignment can be done via handshake mechanisms with the sending host in the Internet. This handshaking would involve the sending host assisting the device in assigning downlink flows to the network suggested by the middleware.

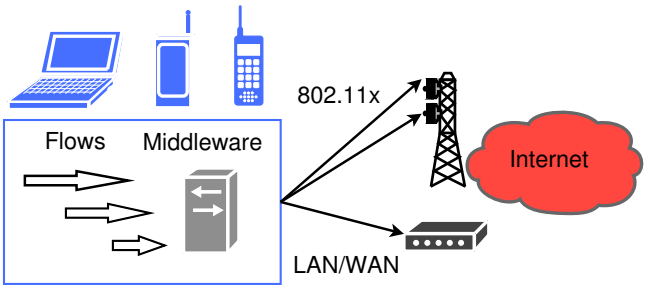

Fig. 1. The envisioned middleware architecture utilizing a lightweight tool to estimate the bandwidth and delay of the heterogeneous networks accessible by the mobile device.

Each device runs a flow control scheme on each network which takes its own measurements of the available bandwidth and delay characteristics as input. Hence, the overall rate control algorithm is fully distributed both in terms of execution and information flow within the system. The objectives of the flow control algorithm are full utilization of the available bandwidth on the access networks, robustness with respect to ABR variations, and prevention of excessive rate uctuations leading to instabilities and jitter. In addition, some applications or flows may have a preference for a particular network due to its favorable delay characteristics.

In order to achieve the performance goals identified above under varying access network characteristics, we resort to a

dynamic and robust algorithm. The H^∞ optimal control theory provides a powerful and widely applicable framework that allows for a worst-case analysis of disturbance attenuation problems. It also allows in this case for the coupling between different devices to be ignored thus leading to a distributed scheme. Using a linear model of the available bandwidth, which can also be interpreted as a low-pass filter, the problem of bandwidth utilization is converted to one of disturbance rejection which is addressed by H^∞-optimal control using a linear feedback control mechanism. A summary of the algorithm is provided in Figure 2 later in the paper.

An interesting question that arises in this setting is the characterization of the access networks in terms of bandwidth and delay. We conduct network measurements in a real world setting and will use the traces to simulate (in Section 7) the flow control schemes introduced in subsequent sections. We monitor the ABR and round trip time (RTT) on different networks including Ethernet and IEEE 802.11b and IEEE 802.11g WLANs. The tests are conducted between hosts in Deutsche Telekom Laboratories (T-Labs) in Berlin to three destinations - Stanford University, Technical University of Munich (TU Munich), and the Technical University of Berlin (TU Berlin) - respectively representing long, mid, and close distance destinations [14]. We surveyed several publicly available tools including Pathrate, Nettest, and CapProbe, and chose Abing [15] for measurement of ABR and RTT.

Table 1 shows the average ABR and RTT and their standard deviations to the TU Berlin destination for different networks for 2 hour traces. We observe that all the networks display noticeable variation in ABRs. For instance the ABR on 802.11g can be as high as 24 Mbps and can drop down to as low as 6 Mbps. These random fluctuations in the ABR observed experimentally clearly indicate the need for a robust flow control approach.

TABLE 1
Available bit rate and RTT from T-Labs to TU Berlin

		ABR(Mbps)	RTT(ms)
Ethernet	Avg.	71.8	5.2
	Std. Dev.	13.0	0.04
802.11g	Avg.	14.3	7.8
	Std. Dev.	3.6	0.4
802.11b	Avg.	4.5	10.7
	Std. Dev.	0.5	0.6

3 Network Model

In this section we present an analytical model of the heterogeneous network (access) environment. We consider a set of access networks, $\mathcal{I} := \{1, 2, \ldots, I\}$, simultaneously available to multiple devices. Let us denote the set of such devices sharing these networks as $\mathcal{D} := \{1, 2, \ldots, D\}$. The assignment and control of flows originating from these devices on these access networks constitute the underlying resource allocation problem we investigate in this paper. Let the nonnegative flow rate a device $d \in \mathcal{D}$ assigns to an available network $i \in \mathcal{I}$ be $r_d^{(i)}$, and let $\mathbf{r}_d := [r_d^{(1)}, r_d^{(2)}, \ldots, r_d^{(I)}]$. Consequently, the total flow rate on network i is $\bar{r}^{(i)} = \sum_{d=1}^{D} r_d^{(i)}$. These quantities will generally be time varying, and hence occasionally we will attach an argument, t, to them to capture the dependence on time. In fact, an important property of the access networks investigated in this paper is the high variability of the network capacity, $C^{(i)}(t)$, where we depict explicitly its dependence on time, t. The available bandwidth $B^{(i)}$ on a network at time t is then given by $B^{(i)}(t) := C^{(i)}(t) - \sum_{d=1}^{D} r_d^{(i)}(t)$. A device can estimate via various online measurement tools [15] the quantity $w^{(i)} := \phi\left(B^{(i)}(t)\right)$, where the function $\phi(\cdot)$ is approximately proportional to its argument, i.e. the available bandwidth $B^{(i)}$.

In our recent study [14], we have investigated a Markov chain based framework for modeling the access network properties and addressed mainly the problem of discrete flows assignment. In this paper we shift our focus to explicit flow control. Towards this end, we introduce and investigate in the next section a linear system formulation and an H^∞ controller that optimizes the network usage.

4 Robust Flow Control

Most of the access networks available to a device at a given time are wireless ones. The characteristics of a (wireless) network $i \in \mathcal{I}$, and hence, its available bandwidth $B^{(i)}$, fluctuate randomly due to fading effects in the case of wireless networks as well as background traffic. Following a different path than the one in [14] we do not attempt to model this quantity but take a function of it $w^{(i)}(B^{(i)})$ simply as an input to devices. We do not make any assumptions on the nature of this function which captures the random variations in available bandwidth due to channel state and other factors.

We next define a system from the perspective of a device $d \in \mathcal{D}$ which keeps track of the available bandwidth of a single access network. The system state $x_d^{(i)}$ reflects from the perspective of device d roughly the bandwidth availability on network i. In order to simplify the analysis we focus in this section on the single network case and drop the superscript i for notational convenience. Then, the system equation for device d is

$$\dot{x}_d = a\, x_d + b\, u_d + w, \tag{1}$$

where u_d represents the control action of the device. The parameters $a < 0$ and $b < 0$ adjust the memory horizon (the smaller a the longer the memory) and the "expected" effectiveness of control actions, respectively, on the system state x_d. The device d bases its control actions on its state which not only takes as input the current available bandwidth but also accumulates the past ones to some extent. It is also possible to interpret the system (1) as a low pass filter with input w and output x.

Let us introduce a rate update scheme which is approximately proportional to the control actions:

$$\dot{r}_d = -\phi r_d + u_d, \tag{2}$$

where $\phi > 0$ is sufficiently small. Although this rate update scheme seems disconnected from the system in (1) it is not the case as we show in Section 4.1. As a result of w being

a function of the available bandwidth B, which in turn is a function of the aggregate user rates, the systems (1) and (2) are connected via a feedback loop.

For simplicity, the coefficient of u_d is chosen to be 1 in (2). Since a rate update of a device will have the inverse effect on the available bandwidth, the parameter b in (1) is naturally picked to be then negative. We additionally note that we resort here to a "bandwidth probing scheme" in a sense similar to additive-increase multiplicative-decrease (AIMD) feature of the well-known transfer control protocol (TCP). On the other hand, our scheme's main parameters follow from an optimization problem which will be defined next.

We now address the question of how to calculate the rate control action u_d of a device. To be able to do so, we first need to formulate our objectives based on the full bandwidth-utilization criterion. We make the following observations on the system described by (1) and (2): First, the input w is zero if all devices fully use the available bandwidth on the access network. Second, the system is stable, i.e., the state converges to zero unless w and u_d are nonzero. Third, the rate change is a function of control actions u_d. Finally, the control actions u_d have a direct effect on the state x_d. Thus, we can formulate the objectives of the optimal controller as minimizing (squares of) w, the state x, and the control actions u. These objectives ensure that the input or "disturbance" to the system is rejected (maximum capacity usage) while preventing excessive rate fluctuations leading to instabilities and jitter.

4.1 Equilibrium and Stability Analysis for Static Capacity

In order to compute the control actions u given the state x, let us consider a linear feedback control scheme of the general form $u = \theta x$, where θ is a positive constant. We conduct an equilibrium and stability analysis of the system (1) and (2) under this general class of linear feedback controllers for a single network of fixed capacity C and accessible by D devices. The analysis of this special no-noise fixed-capacity case provides valuable insights to the original problem.

By ignoring the noise in the system, we make the simplifying assumption of $w := C - \sum_{i=1}^{D} r_i$ and let $d = 1$. Then,

$$\begin{aligned} \dot{x}_i &= a\, x_i + b\, \theta x_i + C - \textstyle\sum_{k=1}^{D} r_k \\ \dot{r}_i &= -\phi r_i + \theta x_i, \quad i = 1, \ldots, D. \end{aligned} \tag{3}$$

At the equilibrium -which we will show below exists, is unique, and is asymptotically stable- we have $\dot{x}_i = \dot{r}_i = 0\ \forall i$. Solving for equilibrium values of x_i and r_i for all i, denoted by x_i^* and r_i^*, respectively, it is easy to obtain

$$r_i^* = \frac{C\theta}{\theta D - (a + b\theta)\phi}$$

and

$$x_i^* = \frac{C\phi}{\theta D - (a + b\theta)\phi}\,,$$

which are unique, under the negativity of a and b and positivity of θ, as long as $\phi > 0$. Now, as $\phi \to 0^+$, we have $\sum_i r_i \to C$. Thus, as ϕ approaches to zero from the positive side, linear feedback controllers of the form $u = \theta x$, where $\theta > 0$ ensure maximum network usage when the capacity C is fixed and there is no noise. We note that the equilibrium rate r^* is on the order of C/D and usually much larger than zero, which constitutes a physical boundary due to nonnegativity constraint. Furthermore, most of the rate control schemes include in practice a *slow start* mechanism that moves r away from zero. Hence, we ignore this constraint in the subsequent stability analysis; an assumption which is also confirmed numerically through simulations in Section 6.

We now show that the linear system (3) is stable and asymptotically converges to the equilibrium point whenever $\phi > 0$. Toward this end, let us sum the rates r_i in (3) to obtain

$$\begin{aligned} \dot{x}_i &= -\mu x_i - \bar{r} + C, \quad i = 1, \ldots, D \\ \dot{\bar{r}} &= -\phi \bar{r} + \theta \textstyle\sum_{i=1}^{D} x_i\ , \end{aligned} \tag{4}$$

where $\bar{r} := \sum_{i=1}^{D} r_i$ and $\mu := -(a + b\theta) > 0$. We can rewrite (4) in the matrix form as

$$\dot{y} = Fy + [C \cdots C\ 0]^T,$$

where $y := [x_1\ \cdots\ x_D\ \bar{r}]^T$. Then, it is straightforward to show that the characteristic function of the $(D + 1)$ dimensional square matrix F has the form

$$\det(sI - F) = (s + \mu)^{D-1} \left[s^2 + (\mu + \phi)s + \mu\phi + D\,\theta\right] = 0.$$

Notice that, F has $D - 1$ repeated negative eigenvalues at $s = -\mu$ and two additional eigenvalues at

$$s_{1,2} = -\frac{1}{2}(\mu + \phi) \pm \frac{1}{2}\sqrt{(\mu - \phi)^2 - 4D\,\theta}.$$

If $(\mu - \phi)^2 < 4D\,\theta$, then both of these eigenvalues are imaginary with negative real parts. Otherwise, we have $\mu + \phi > |\mu - \phi|$ and both eigenvalues are negative and real. Therefore, all eigenvalues of F always have negative real parts and the linear system (4) is stable.

It immediately follows that x_i is always finite and converges to the equilibrium, and from the second equation of (3), r_i has to be finite and converges for all i. We thus establish that the original system (3) is stable.

4.2 H$^\infty$-Optimal Control

Having obtained the equilibrium state of (3) and shown its asymptotic stability, we now turn to robustness analysis. We first rewrite the system (3) around the equilibrium point $(x_i^*, r_i^*)\ \forall i$ to obtain

$$\begin{aligned} \dot{\tilde{x}}_i &= a\, \tilde{x}_i + b\, \tilde{u}_i + w \\ \dot{\tilde{r}}_i &= -\phi \tilde{r}_i + \theta \tilde{x}_i, \quad i = 1, \ldots, D, \end{aligned} \tag{5}$$

where $\tilde{x}_i := x_i - x_i^*$, $\tilde{r}_i := r_i - r_i^*$, and $\tilde{u}_i := u_i - u_i^*$. We then reformulate the objectives described earlier within a disturbance rejection problem around the equilibrium. Subsequently, H$^\infty$ optimal control theory allows us to remove all the simplifying assumptions of the previous subsection on w and solve the problem in the most general case as described earlier. By viewing the disturbance (here the available bandwidth) as an intelligent maximizing opponent in a dynamic zero-sum game

who plays with knowledge of the minimizer's control action, we evaluate the system under the worst possible conditions (in terms of capacity usage). We then determine the control action that will minimize costs or achieve the objectives defined under these worst circumstances [16]. Hence, we obtain a robust linear feedback rate control scheme.

Notice that, we assume here a time scale separation between the variations in capacity $C(t)$ and the rate updates $r(t)$. With a sufficiently high update frequency each device can track the variations at the equilibrium point caused by the random capacity fluctuations. The robustness properties of H^∞ optimal controller also play a positive role here. We verify this assumption numerically through simulations in Section 6.

The system (1) can be classified as continuous-time with perfect state measurements due to the state $\tilde{x}_d$ being an internal variable of device d. We conduct an H^∞-optimal control analysis and design taking this into account. Let us first introduce the *controlled output*, $z_d(t)$, as a two dimensional vector:

$$z_d(t) := [h\tilde{x}_d(t) \;\; g\tilde{u}_d(t)]^T, \tag{6}$$

where g and h are positive (network-specific) parameters. The cost of a device that captures the objectives defined and for the purpose of H^∞ analysis is the ratio of the L^2-norm of z_d to that of w:

$$L_d(\tilde{x}_d, \tilde{u}_d, w) = \frac{\|z_d\|}{\|w\|}, \tag{7}$$

where $\|z_d\|^2 := \int_0^\infty |z_d|^2 \, d\tau$, and a similar definition applies to $\|w\|^2$. Although being a ratio, we will refer to L_d as the (device) cost in the rest of the analysis. It captures the proportional changes in z_d due to changes in w. If $\|w\|$ is very large, the cost L_d should be low even if $\|z_d\|$ is large as well. A large $\|z_d\|$ indicates that the state $|\tilde{x}_d|$ and/or the control $|u_d|$ have high values reflecting and reacting to the situation, respectively. However, they should not grow unbounded, which is ensured by a low cost, L_d. For the rest of the analysis, we will drop the subscript d for ease of notation.

H^∞-optimal control theory guarantees that a performance factor will be met. This factor γ, also known as the H^∞ norm, can be thought of as the worst possible value for the cost L. It is bounded below by

$$\gamma^* := \inf_{\tilde{u}} \sup_{w} L(\tilde{u}, w), \tag{8}$$

which is the lowest possible value for the parameter γ. It can also be interpreted as the optimal performance level in this H^∞ context. Interestingly, we assume here that the available bandwidth is "controlled" by a maximizing player (we call as Murphy) who plays second in this formulation knowing the control applied by the minimizing player or device. This formulation as well as the order of play ensures that we are indeed analyzing the worst case scenario.

In order to solve for the optimal controller $\mu(\tilde{x})$, a corresponding (soft-constrained) differential game is defined, which is parametrized by γ,

$$J_\gamma(\tilde{u}, w) = \|z\|^2 - \gamma^2 \|w\|^2. \tag{9}$$

The maximizing player (Murphy) tries to maximize this cost function while the objective of a device is to minimize it. The optimal control action $\tilde{u} = \mu_\gamma(\tilde{x})$ can be determined from this differential game formulation for any $\gamma > \gamma^*$.

This controller is expressed in terms of a relevant solution, σ_γ, of a related game algebraic Riccati equation (GARE) [16]:

$$2a\sigma - \left(\frac{b^2}{g^2} - \frac{1}{\gamma^2}\right)\sigma^2 + h^2 = 0 \tag{10}$$

By the general theory [12], the relevant solution of the GARE is the "minimal" one among its multiple nonnegative-definite solutions. However, in this case, since the GARE is scalar, and the system is open-loop stable (that is, $a < 0$), the GARE (which is a quadratic equation) admits a unique positive solution for all $\gamma > \gamma^*$, and the value of γ^* can be computed explicitly in terms of the other parameters. Solving for the roots of (7), we have:

$$\sigma_\gamma = \frac{-a \pm \sqrt{a^2 - \lambda h^2}}{\lambda}$$

where

$$\lambda := \frac{1}{\gamma^2} - \frac{b^2}{g^2}.$$

The parameter λ could be both positive and negative, depending on the value of γ, but for γ close in value to γ^* it will be positive. Further, γ^* is the smallest value of γ for which the GARE has a real solution. Hence,

$$\gamma^* = \left[\sqrt{\frac{a^2}{h^2} + \frac{b^2}{g^2}}\right]^{-1}$$

Finally, a controller that guarantees a given performance bound $\gamma > \gamma^*$ is:

$$u_\gamma = \mu_\gamma(x) = -\left(\frac{b}{g^2}\,\sigma_\gamma\right)\, x. \tag{11}$$

This is a stabilizing linear feedback controller operating on the device system state x, where the gain can be calculated offline using only the linear quadratic system model and for the given system and cost parameters.

It is important to note that although we conduct the analysis and controller design around the equilibrium point, the devices do not have to compute the actual equilibrium values. In other words, (11) can be equivalently written in terms of $\tilde{u}_\gamma$ and $\tilde{x}$. In practice, the H^∞-optimal rate control scheme is implemented as follows: each device keeps track of the ABR of one or more access networks via the respective state equation (1), which takes the respective w as input. The linear feedback control u is computed in (11) for each network separately for a given set of system (a, b) and preference (h, g) parameters. Finally, the device updates its flow rate using (2) on each network. A discretized version of the algorithm is summarized in Figure 2.

4.3 Alternative Controller Formulations

Having established and analyzed the H^∞-optimal controller for the system at hand, we study variations of it and other formulations. One possible formulation is the well-known Linear-Quadratic-Gaussian (LQG) problem where the input w is modeled as a Gaussian noise. Although this assumption probably does not hold for the problem at hand we use the

Input: ABR and delay measurements of available access networks;
Parameters: Device (a, b) and network preference (h, g) for each network;
Output: Feedback control u for each device and rate r for each network;

foreach *Access network available to Device* **do**
  Measure current ABR (w) and delay ;
  Update x using (1) ;
  Compute u using (11) ;
  Update flow rate r using (2) ;
end

Fig. 2. H$^\infty$-optimal rate control scheme run by a device

LQG model as a comparison case. It can be obtained here simply as the limit of the H$^\infty$ control problem as $\gamma \to \infty$. We use as the $\|z\|^2$, the expected value of $\int_0^\infty |z|^2\, dt$, which we again denote by $\|z\|^2$ by a slight abuse of notation, and the problem is one of minimization of $\|z\|^2$.

As an alternative, we define a simple AIMD controller as another comparison scheme:

$$\dot{r}_d = \begin{cases} \alpha & ,\text{if } w > 0 \\ -\beta r & ,\text{if } w < 0 \end{cases}, \tag{12}$$

where α and β are positive parameters.

4.4 An Illustrative Example

We illustrate the H$^\infty$-optimal controller with an example. The cost and system parameters are chosen simply as $a = -1$, $b = -1$, $g = 1$, $h = 1$. Then, $\gamma^* = \sqrt{2}/2 \approx 0.707$. If we choose $\gamma = \gamma^*$, then the unique positive solution of the GARE is $\bar{\sigma}_{\gamma^*} = 1$ which leads to the simple feedback controller $\mu_{\gamma^*}(x) = x$, which is the optimal H$^\infty$ controller.

We compare this result analytically with the LQG formulation where $\gamma \to \infty$. Then, the GARE, $\sigma^2 + 2\sigma - 1 = 0$, simply yields the unique positive solution $\tilde{\sigma}_\gamma \approx \sqrt{2} - 1$, which leads to $\tilde{\mu}_\gamma(x) \approx (\sqrt{2} - 1)x$. We observe that despite the same cost structure, the optimal H$^\infty$ controller is more "aggressive" in order to ensure an upper bound on the cost L regardless of w. On the other hand, the LQG controller has a lower feedback gain possibly due to the inherent Gaussian noise assumption on w.

5 H$^\infty$-optimal control for Multiple Networks

In Section 4, we have provided analysis and controller design for a *single* access network shared by multiple devices, mainly to enhance readability and focus on core concepts by keeping the notation simple. We now provide the H$^\infty$-optimal control formulation for the general case of multiple access networks for each device $d \in \mathcal{D}$, and drop the subscript d again for ease of notation.

Assuming a fixed capacity vector, let us introduce $\tilde{\mathbf{x}} := [\tilde{x}^{(i)}]$, $\tilde{\mathbf{r}} := [\tilde{r}^{(i)}]$, and $\tilde{\mathbf{u}} := [\tilde{u}^{(i)}]$ as I-dimensional vectors around the respective equilibria, with components defined for $i \in \mathcal{I}$. Then, the counterpart of the system (5) is given by

$$\begin{aligned} \dot{\tilde{\mathbf{x}}} &= A\,\tilde{\mathbf{x}} + B\,\tilde{\mathbf{u}} + D\,\mathbf{w} \\ \dot{\tilde{\mathbf{r}}} &= -\Phi\,\tilde{\mathbf{r}} + \tilde{\mathbf{u}}, \end{aligned} \tag{13}$$

where $\mathbf{w} := [w^{(i)}] \;\; \forall i$. Here, the matrices A, B, and Φ are obtained simply by multiplying the identity matrix by a, b, and ϕ, respectively.

The counterpart of the *controlled output* in (6) is

$$\mathbf{z}(t) := H\tilde{\mathbf{x}}(t) + G\tilde{\mathbf{u}}(t), \tag{14}$$

where we assume that $G^T G$ is positive definite, and that no cost is placed on the product of control actions and states, *i.e.,* $H^T G = 0$. The matrix H represents a cost on variation from zero state, i.e. full capacity usage, and we make the natural assumption that $Q := H^T H$ is positive definite.

We next define the cost

$$L(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \mathbf{w}) = \frac{\|\mathbf{z}\|}{\|\mathbf{w}\|}, \tag{15}$$

where $\|\mathbf{z}\|^2 := \int_0^\infty |\mathbf{z}(t)|^2 dt$, and the corresponding differential game parametrized by γ,

$$J_\gamma(\tilde{\mathbf{u}}, \mathbf{w}) = \|\mathbf{z}\|^2 - \gamma^2 \|\mathbf{w}\|^2 \tag{16}$$

as in Section 4.2. Here γ is larger than γ^*, where γ^* is defined as in (8).

The corresponding GARE

$$A^T Z + ZA - Z(B(G^T G)^{-1} B^T - \gamma^{-2} DD^T)Z + Q = 0, \tag{17}$$

admits a unique minimal nonnegative definite solution $\bar{Z}_\gamma$, for $\gamma > \gamma^*$, if (A, B) is stabilizable and (A, H) is detectable [16]. In our case, since B is square and negative definite, and Q is positive definite, the system is both controllable and observable, and hence both conditions are satisfied. Thus, we obtain the H$^\infty$-optimal linear feedback controller for the multiple network case:

$$\mu_\gamma(\tilde{\mathbf{x}}) = -(G^T G)^{-1} B^T \bar{Z}_\gamma \tilde{\mathbf{x}}, \tag{18}$$

for each $\gamma > \gamma^*$, which is also stabilizing. We finally note that the same feedback controller above can be written in terms of $\mathbf{x}$ instead of $\tilde{\mathbf{x}}$. Again, the devices do not have to compute the equilibrium values in order to follow the algorithm in Figure 2.

6 Simulations

We simulate the H$^\infty$-optimal controller in a scenario where 20 devices share three different network interfaces with varying available bandwidths obtained from real world measurements as discussed in Section 2. We first study the illustrative case where the bandwidth available on a given network is the same for all devices. We then investigate a more realistic scenario in Section 6.1 where devices make noisy and delayed measurements. These results play an indicative role for settings where individual devices have different available bandwidths due to, for example, variations in individual channel gains.

The system parameters are $a = -1$, $b = -1$, $g = 1$ and chosen to be the same over all three networks resulting in G to be the identity matrix and A and B to be the negative identity matrices. For the last 15 devices, the Q matrix is chosen as the identity matrix while for the first 5 devices it is

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which indicates a preference for network 2 due to, for example, favorable delay characteristics and nature of applications running on these devices. Hence, the controllers $u1$ and $u2$ for the first 5 and last 15 devices, respectively, are

$$u1 = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 3.9 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \cdot \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ x^{(3)} \end{bmatrix},$$

$$u2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ x^{(3)} \end{bmatrix}.$$

The corresponding values of γ^* are calculated as 0.895 and 0.707, respectively.

The available bitrate (ABR) is -by definition- always positive, $B(t) > 0$. Ideally, the algorithm needs the information on the exact difference between the aggregate rate and the capacity, $C(t) - B(t)$, even when the value is negative. However, the measurement system does not provide this information in real life. Hence, we choose

$$w = \begin{cases} B(t) & , \text{if } B(t) > 0 \\ -0.1\, C(t) & , \text{if } B(t) = 0 \end{cases},$$

where the negative value in the case when the available bandwidth hits zero, ensures a proper reaction of the controller. Here, we take into account the fact that the aggregate flow rate will surpass the capacity given the trend.

Each device updates its flows according to the algorithm in Figure 2. The resulting aggregate flow rates and capacity of each network are depicted in Figures 3, 4, and 5. The average network usage on the networks is approximately 89 to 90%. The corresponding individual flows of devices on each network are shown in Figures 6, 7, and 8. As expected, the 5 devices with a preference for network 2 receive a higher share of bandwidth on it.

We next compare the H$^\infty$-optimal controller with the AIMD scheme in (12) with parameters $\alpha = 10$ and $\beta = 0.75$ for 20 symmetric devices on network 1. The results are depicted in Figure 9. We observe that the average capacity usage under the AIMD controller is only 75% and the H$^\infty$-optimal controller outperforms AIMD in this respect. In addition, the flow rates fluctuate less under the H$^\infty$ scheme despite a careful choice of AIMD parameters.

As a second comparison, we simulate the LQG controller discussed in 4.4 within the same environment. As shown in Figure 10 the LQG performs better than the AIMD but worse than the H$^\infty$-optimal controller with an average capacity usage of 85%. We observe that the aggregate flow rate does not

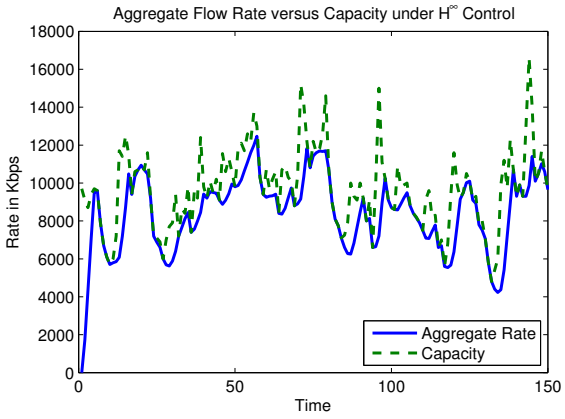

Fig. 3. The aggregate flow rate and the available capacity on network 1 under H$^\infty$-optimal control.

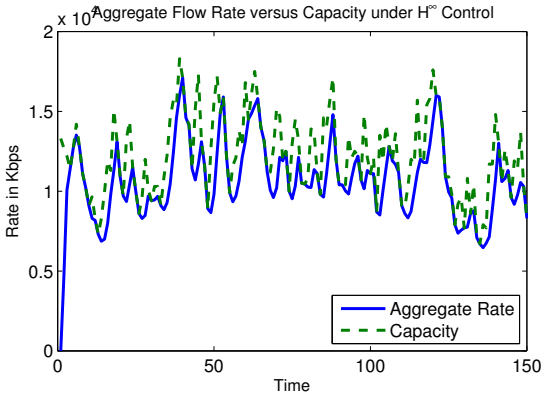

Fig. 4. The aggregate flow rate and the available capacity on network 2 under H$^\infty$-optimal control.

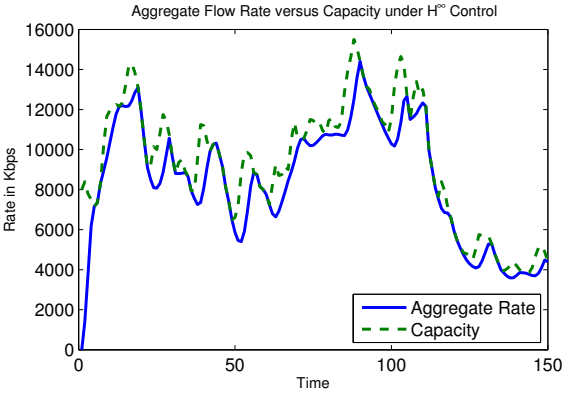

Fig. 5. The aggregate flow rate and the available capacity on network 3 under H$^\infty$-optimal control.

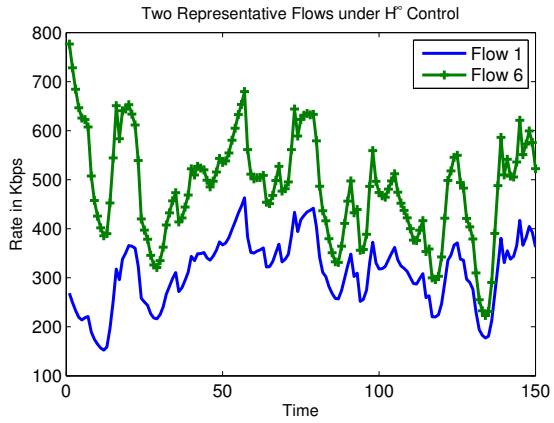

Fig. 6. The rates of representative flows 1 and 6 on network 1 under H^{∞}-optimal control.

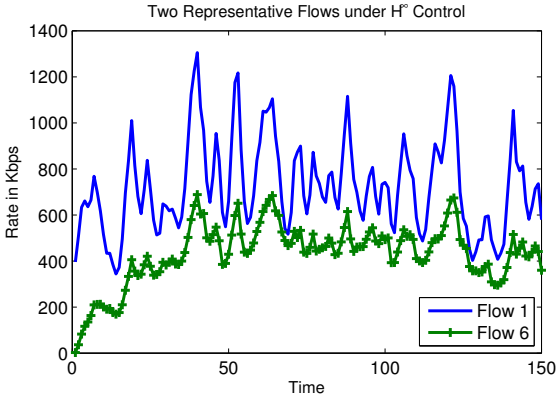

Fig. 7. The rates of representative flows 1 and 6 on network 2 under H^{∞}-optimal control.

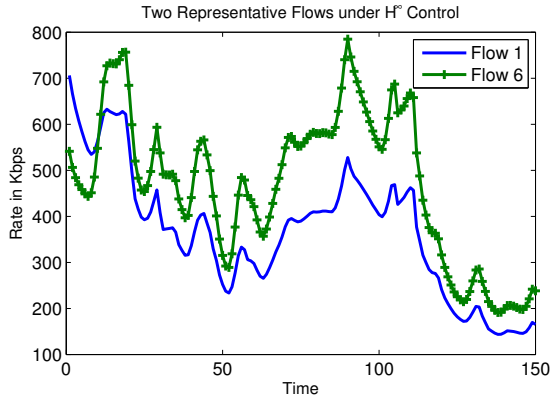

Fig. 8. The rates of representative flows 1 and 6 on network 3 under H^{∞}-optimal control.

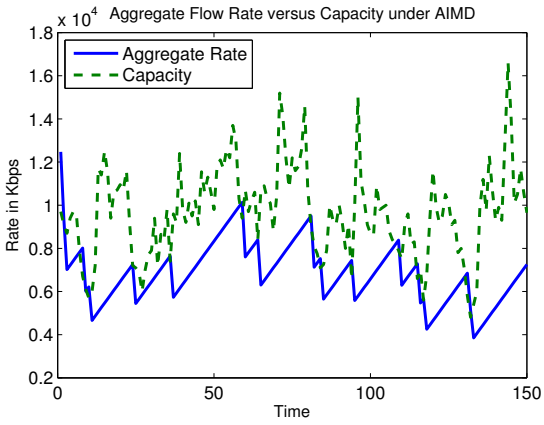

Fig. 9. The aggregate flow rate and the available capacity on network 1 under AIMD control.

follow the capacity as closely as it was the case with the H^{∞}-optimal control. It is also important to note that the LQG scheme does not provide a minimum performance guarantee on the cost L whereas the H^{∞} control ensures one.

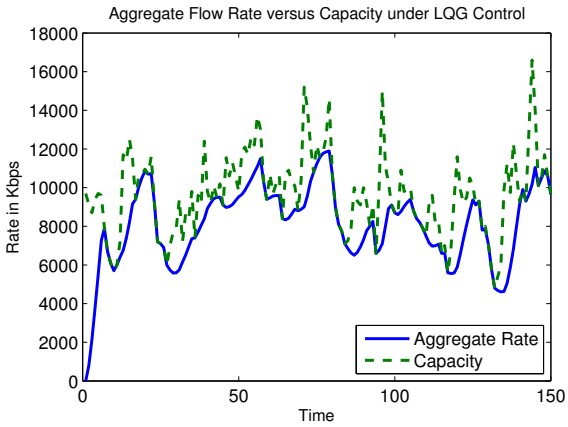

Fig. 10. The aggregate flow rate and the available capacity on network 1 under LQG control.

Subsequently, we investigate the fairness properties of our approach by simulating 20 devices with random initial flow rates on network 1. We observe in Figure 11 that within a short time each flow converges to an equal share of the available bandwidth or capacity. Previous simulation results in Figures 6, 7, and 8 also show that the devices which obtain more bandwidth on network 2 get less on the other two networks further indicating the fairness of our approach.

6.1 Robustness Analysis

We study the robustness properties of H^{∞}-optimal controller with respect to measurement noise and delays as well as variations in system parameters and the number of devices. In the first simulation, we abruptly increase the number of devices accessing the network from 20 to 40 at time step $t = 100$ and deactivate them again at $t = 200$. We observe in Figures 12 and 13 that our algorithm successfully responds to these changes with a high speed of convergence.

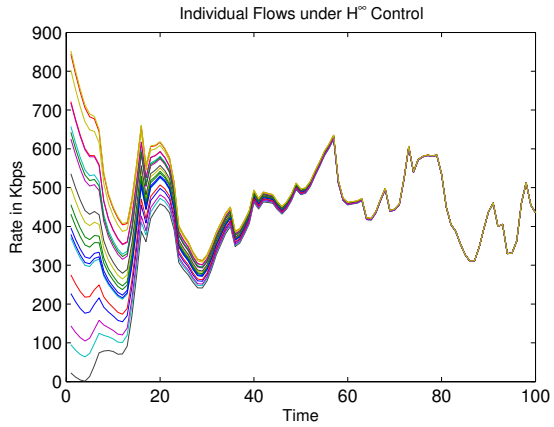

Fig. 11. The rates of individual flows with random starting points on network 1 under H^∞-optimal control.

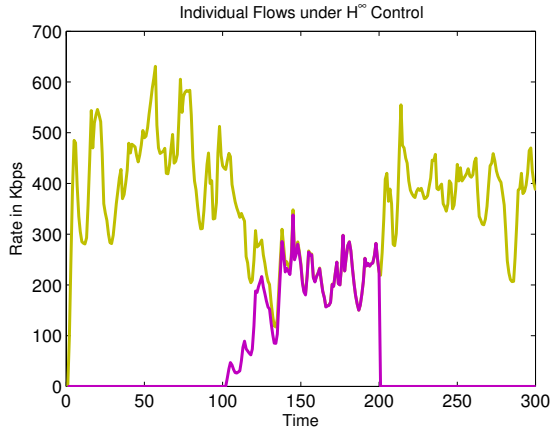

Fig. 12. The flow rates on network 1 under H^∞-optimal control and varying number of active devices.

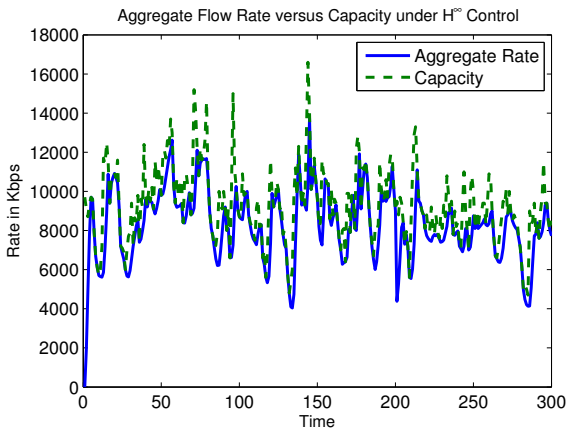

Fig. 13. The aggregate flow rate and the available capacity on network 1 under H^∞-optimal control and varying number of active devices.

Next, we investigate the effect of parameter b on capacity usage in the system. A high value of the parameter b means that the actions of the device have a significant effect on the network while a smaller value corresponds to the case where a large number of devices share the capacity. We observe in Table 2 that the algorithm is fairly robust with respect to variations in parameter b while the number of devices is kept constant at 20.

TABLE 2
Average capacity usage versus parameter b

b	0.1	1	2	5	10
Capacity Usage (%)	90	90	89	86	80

We introduce a Gaussian noise of zero mean and standard deviation 200 to the ABR measurements, i.e. $w = B + \mathcal{N}(0, 200^2)$, if $B > 0$. We observe in Figures 14 and 15 that our H^∞-optimal control successfully rejects the measurement noise and maintains high bandwidth utilization. Finally, we consider the case where devices operate under

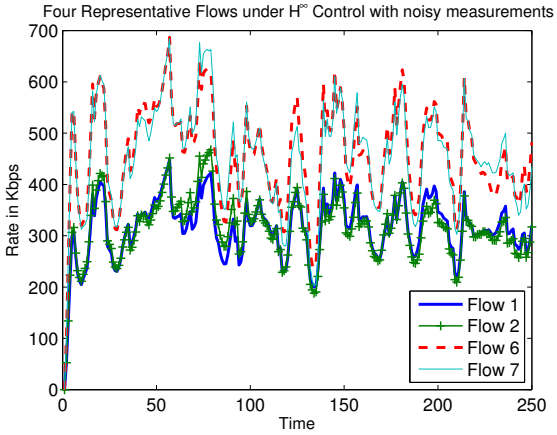

Fig. 14. The flow rates of individual flows on network 1 under H^∞-optimal control and Gaussian measurement noise $\mathcal{N}(0, 200^2)$.

randomly delayed measurements: $w = B(t-\tau)$, where τ takes uniformly random delay values. As an example, we consider a system that is in accordance with the measurements in Table 1 where the flows $1, 2, 6, 7$ experience individual measurement delays of $2.5, 10, 7.5, 12.5$ ms, respectively. We observe higher variance at the individual flow level under measurement delays (see Figure 16), when compared to the ones without delays (Figure 14). However, the algorithm is observed to be robust with respect to aggregate bandwidth utilization (Figure 17). The bandwidth utilization decreases approximately 2% in both cases.

7 Conclusion

We have presented a robust flow control approach based on H^∞-optimal control theory for the purpose of efficient utilization of multiple heterogeneous networks and a fair bandwidth

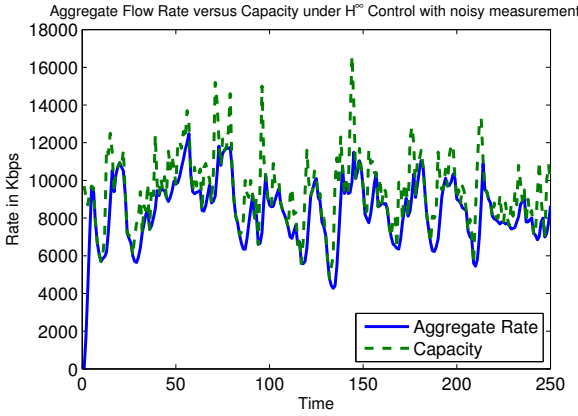

Fig. 15. The aggregate flow rate and the available capacity on network 1 under H^∞-optimal control and Gaussian measurement noise $\mathcal{N}(0, 200^2)$.

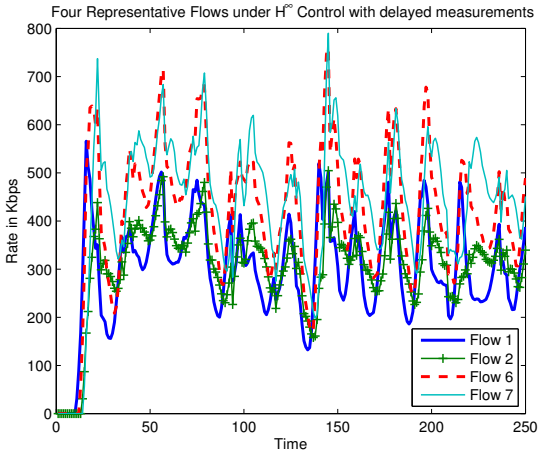

Fig. 16. The flow rates of individual flows on network 1 under H^∞-optimal control. In accordance with Table 1, flows $1, 2, 6, 7$ experience individual measurement delays of $2.5, 10, 7.5, 12.5$ ms, respectively.

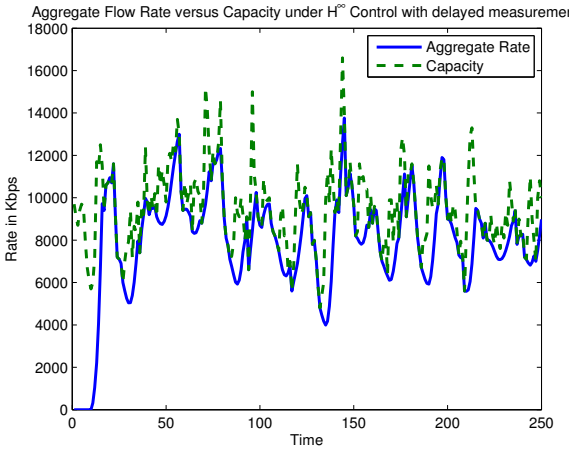

Fig. 17. The aggregate flow rate and the available capacity on network 1 under H^∞-optimal control and randomly delayed measurements.

allocation to devices accessing them. Bandwidth and delay measurements of different network types in a real world setting have indicated random fluctuations of these quantities and justified the necessity of a robust rate control scheme. We have modeled the system from a device's perspective and derived a minimum information rate control scheme using optimal control actions obtained through H^∞ analysis and design. By reformulating the rate control problem as one of disturbance rejection, we have utilized H^∞ control theory without making any restrictive assumptions on the random nature of network characteristics.

An efficient utilization of the access networks under our algorithm has been established through an equilibrium analysis in the static case. We have considered an LQG (as a variation of the H^∞) control scheme as well as a simple AIMD algorithm for comparison purposes. The efficiency, fairness, and robustness properties of the H^∞-optimal rate controller developed have been demonstrated via simulations using the measured real world network characteristics.

The promising results obtained are motivating for future research. One immediate direction for extension is to take the x_i and r_i dynamics to be device dependent, that is for the parameters a, b, ϕ and θ to be indexed by i. The equilibrium values of x_i and r_i can also be readily computed in that case, but the proof of asymptotic stability seems to be fairly involved, though tractable. Finally, we note that the robust flow control scheme can also be applied to wired networks.

REFERENCES

[1] H. Yaiche, R. Mazumdar, and C. Rosenburg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," in *IEEE/ACM Transaction on Networking*, vol. 8, no. 5, Oct. 2000, pp. 667–678.

[2] T. Alpcan and T. Başar, "A utility-based congestion control scheme for Internet-style networks with delay," *IEEE Transactions on Networking*, vol. 13, no. 6, pp. 1261–1274, December 2005.

[3] ——, "Global stability analysis of an end-to-end congestion control scheme for general topology networks with delay," in *Proc. of the 42nd IEEE Conference on Decision and Control*, Maui, HI, December 2003, pp. 1092 – 1097.

[4] K. Chebrolu and R. Rao, "Communication using multiple wireless interfaces," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC 2002)*, vol. 1, 2002, pp. 327–331.

[5] P. Vidales, J. Baliosion, J. Serrat, G. Mapp, F. Stejano, and A. Hopper, "Autonomic sytem for mobility support in 4G networks," *IEEE JSAC*, vol. 23, no. 12, Dec. 2005.

[6] A. Zemlianov and G. de Veciana, "Cooperation and decision-making in a wireless multi-provider setting," in *Proc. IEEE INFOCOM*, 2005, pp. 1–14.

[7] E. Altman and T. Başar, "Optimal rate control for high speed telecommunication networks," in *Proc. of the 34th IEEE Conference on Decision and Control*, New Orleans, LA, December 1995, pp. 1389–1394.

[8] N. Thompson, G. He, and H. Luo, "Flow scheduling for end-host multihoming," in *Proc. IEEE INFOCOM*, 2006.

[9] S. Shakkottai, E. Altman, and A. Kumar, "Multihoming of users to access points in WLANs: A population game perspective," *IEEE JSAC*, vol. 25, no. 6, pp. 1207–1215, August 2007.

[10] D. Kumar, E. Altman, and J.-M. Kelif, "Globally optimal user-network association in an 802.11 WLAN and 3G UMTS hybrid cell," in *Proc. of 20th Int. Teletraffic Congress (ITC-20)*, Ottawa, Canada, June 2007.

[11] R. Chandra and P. Bahl, "Multinet: connecting to multiple IEEE 802.11 networks using a single wireless card," in *Proc. 23rd IEEE Infocom*, vol. 2, Hong Kong, China, March 2004, pp. 882– 893.

[12] X. Zhu, P. Agarwal, J. P. Singh, T. Alpcan, and B. Girod, "Rate allocation for multi-user video streaming over heterogeneous access networks," in *Proc. ACM Multimedia*, Augsburg, Germany, September 2007.

[13] J. P. Singh, T. Alpcan, X. Zhu, and P. Agarwal, "Towards heterogeneous network convergence: policies and middleware architecture for efficient flow assignment, rate allocation, and rate control for multimedia applications," in *Proc. MNCNA 2007: Middleware for Next-generation Converged Networks and Applications, Workshop of the 8th Int. Middleware Conf.*, Newport Beach, CA, November 2007.

[14] J. P. Singh, T. Alpcan, P. Agrawal, and V. Sharma, "An optimal flow assignment framework for heterogeneous network access," in *Proc. IEEE Intl. Symp. on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Helsinki, Finland, June 2007.

[15] J. Navratil and R. L. Cottrell. Abing. [Online]. Available: http://www-iepm.slac.stanford.edu/tools/abing/

[16] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd ed. Boston, MA: Birkhäuser, 1995.

Tansu Alpcan (S'98-M'07) received the B.S. degree in electrical engineering from Bogazici University, Istanbul, Turkey in 1998. He received the M.S. and PhD degrees in electrical and computer engineering from University of Illinois at Urbana-Champaign (UIUC) in 2001 and 2006, respectively. His research interests include network security, game theory, control and optimization of wired and wireless communication networks, resource allocation, and intrusion detection. He has received Fulbright scholarship in 1999 and best student paper award in IEEE Conference on Control Applications in 2003. He is the (co-)author of more than 50 journal and conference articles. He was an associate editor for IEEE Conference on Control Applications (CCA) in 2005 and has been TPC member of several conferences including IEEE Infocom in 2007, 2008, and 2009. He is the co-chair of the workshop on Game Theory in Communication Networks (GameComm) 2008. He has received Robert T. Chien Research Award from the UIUC Department of Electrical and Computer Engineering and Ross J. Martin Research Award from the UIUC College of Engineering in 2006. Tansu Alpcan has been a member of IEEE since 1998. He is currently a senior research scientist in Deutsche Telekom Laboratories, which is affiliated with Technische Universitat Berlin, Germany.

Jatinder Pal Singh is a Consulting Assistant Professor with the department of Electrical Engineering at Stanford University and Director of Research with Deutsche Telekom Laboratories. He was a Senior Research Scientist with the labs from 2006-2007. He received his Ph.D. (2005) and M.S. (2002) in Electrical Engineering from Stanford University, where he was awarded Stanford Graduate Fellowship and Deutsche Telekom Fellowship. He received his B.S. (2000) in Electrical Engineering from the Indian Institute of Technology, Delhi, where he graduated at the top of his class with Institute Silver Medal and A.K. Mahalanabis memorial scholarship. He worked with IBM T.J. Watson Research Center (Hawthorne, NY), Robert Bosch Corporation (Palo Alto, CA), Intel Corporation (San Jose, CA), and IBM India Research Labs (New Delhi) as an intern. He has been named Who's Who in Science and Engineering (2008) by Marquis Who's Who Publishers. His current research focus is on next generation fixed and mobile infrastructure and services, resource allocation for heterogeneous network access, , and media streaming over wireless and peer to peer networks,. His previous research spanned the areas of secure and controlled broadband access sharing, cross-layer design issues in wireless networks, routing in ad-hoc and vehicular networks, performance evaluation of 802.11 based WLANs for inter-vehicle communication, application layer multicast, peer to peer media streaming, and transport layer optimization via adaptive link-layer techniques. He has published in and served as a reviewer for prominent journals and conferences in networking and communications. He has authored patents (granted and pending) in the areas of load balancing, QoS provisioning and distributed authentication mechanisms for broadband access sharing, TCP throughput optimization, central scheduling algorithms for IEEE 802.11e networks, peer-to-peer media streaming, and heterogeneous network access. He is a member of the IEEE and ACM.

Tamer Başar (S'71-M'73-SM'79-F'83) received B.S.E.E. degree from Robert College, Istanbul, and M.S., M.Phil, and Ph.D. degrees in engineering and applied science from Yale University.

After working at Harvard University and Marmara Research Institute (Gebze, Turkey), he joined the University of Illinois at Urbana-Champaign in 1981, where he is currently the Swanlund Endowed Chair in the Department of Electrical and Computer Engineering, and Center for Advanced Study Professor. He has published extensively in systems, control, communications, and dynamic games, and has current interests in modeling and control of communication networks, control over heterogeneous networks, resource management, pricing, and security in networks, and robust identification and control.

Dr. Başar is the Editor-in-Chief of Automatica, Editor-in-Chief of IEEE Expert Now, Editor of the Birkhäuser Series on Systems & Control, Managing Editor of the Annals of the International Society of Dynamic Games (ISDG), and member of editorial and advisory boards of several international journals. He has received several awards and recognitions over the years, among which are the Medal of Science of Turkey (1993), and Distinguished Member Award (1993), Axelby Outstanding Paper Award (1995) and Bode Lecture Prize (2004) of the IEEE Control Systems Society (CSS), the triennial Quazza Medal (2005) of the International Federation of Automatic Control (IFAC), and the Richard E. Bellman Control Heritage Award (2006) of the American Automatic Control Council (AACC). He is a member of the National Academy of Engineering, a member of the European Academy of Sciences, a Fellow of IFAC, president-elect of AACC, a past president of CSS, and the founding president of ISDG.