

Convergence of REM Flow Control at a Single Link

Qinghe Yin and Steven H. Low, *Senior Member, IEEE*

Abstract—Various TCP congestion control schemes can be interpreted as approximately carrying out a certain basic algorithm to maximize aggregate source utility, different schemes corresponding to different choices of utility functions. The basic algorithm consists of a link algorithm that updates a congestion measure based on its traffic load, and a source algorithm that adapts the source rate to congestion in its path. Though convergent, this algorithm can lead to large equilibrium backlogs. This problem can be eliminated by modifying the basic algorithm to include backlog in the update of the congestion measure. This letter proves that the modified algorithm converges when the network can be modeled as a single bottleneck link. Moreover, in equilibrium, the source rates are matched to link capacity while the buffer is cleared.

Index Terms—Active queue management, congestion control, convergence, REM, stability.

I. INTRODUCTION

WE HAVE proposed in [10] a duality model of flow control¹ from which a basic algorithm is derived to maximize aggregate source utility. The significance of the model is that TCP Reno [5], [12] and Vegas [3], with or without random marking [4], [2], can all be interpreted within this model as *approximately* carrying out the basic algorithm to maximize aggregate utility, different schemes merely corresponding to different choices of utility functions; see [9]². The model thus provides a convenient way to understand the stability, optimality and fairness properties of these schemes, and more importantly, to explore their interaction. The basic algorithm consists of a link algorithm that updates a congestion measure, called price, based on its load, and a source algorithm that adapts the source rate to congestion in its path. It is proved in [10] that it converges even in an asynchronous environment provided the time intervals between updates are bounded.

Straightforward implementation of the basic algorithm can lead to large equilibrium backlogs in the network; see Section II. In [2] we correct this problem by including buffer occupancy in price adjustment. The purpose of this paper is to prove that the modified algorithm, called REM, still converges when the network is modeled as a single (bottleneck) link. Moreover, in

equilibrium, the source rates are matched to link capacity while the buffer is cleared, as observed in extensive simulations in [2], [1].

In Section II we review the duality model of flow control, and describe the basic algorithm and REM. The convergence proof for the basic algorithm relies on the fact that the dual objective function (on the prices) serves as a Lyapunov function for the algorithm so that its value monotonically decreases as the algorithm proceeds. The addition of the buffer term in the price adjustment destroys this monotonicity, and a different approach is needed. In Section III we provide a convergence proof for the single-link case.

Our proof does not generalize to the multilink case. However if we model the algorithm as a continuous time system, then it can be shown that it is *globally* asymptotically stable in a multilink network, i.e., starting from any initial value, the rates converge to the unique optimal point; see [11]. A local convergence proof for a discrete time multilink network is presented in [1].

II. PRELIMINARIES

We briefly review the model, the basic algorithm, and REM developed in [10], [2].

A. Single-Link Model

Consider a network that is modeled as a single (bottleneck) link of capacity c . It is shared by a set S of sources. Source s attains a utility of $U_s(x_s)$ when it transmits at rate x_s that satisfies $0 \leq m_s \leq x_s \leq M_s < \infty$. We assume U_s are strictly concave increasing and twice differentiable.

Our objective (the primal problem) is to choose source rates $x = (x_s, s \in S)$ so as to

$$\max_{m_s \leq x_s \leq M_s} \sum_s U_s(x_s) \quad (1)$$

$$\text{subject to } \sum_s x_s \leq c. \quad (2)$$

Constraint (2) says that the aggregate source rate does not exceed the capacity. This flow control problem is posed in [6] and solved in [7], [8] using a penalty function approach. The basic algorithm is a gradient projection algorithm to solve the dual problem.

B. Basic Algorithm

Associated with the link is a (scalar) dual variable p we call *price*. The basic algorithm is given by the following iteration:

$$p(t+1) = \left[p(t) + \gamma \left(\sum_s x_s(t) - c \right) \right]^+ \quad (3)$$

Manuscript received August 31, 2000. The associate editor coordinating the review of this letter and approving it for publication was Dr. I. S. Venieris. This work was supported by the Australian Research Council under Grant S499705, Grant A49930405, and Grant S4005343.

Q. Yin is with the ARC Centre for Ultra-Broadband Information Network, Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, Australia (e-mail: qyin@ee.mu.oz.au).

S. H. Low is with the CS and EE Departments, Caltech, Pasadena, CA 91125 USA (e-mail: slow@caltech.edu).

Publisher Item Identifier S 1089-7798(01)02811-3.

¹We use “flow control” and “congestion control” interchangeably.

²Online. Available: <http://netlab.caltech.edu>

$$x_s(t) = x_s(p(t)), \quad s \in S \quad (4)$$

where $[z]^+ = \max\{0, z\}$ and $x_s(p) = [U'_s{}^{-1}(p)]_{m_s}^{M_s} := \max\{m_s, \min\{M_s, U'_s{}^{-1}(p)\}\}$. Here $U'_s{}^{-1}(\cdot)$ denote the inverse functions of the marginal utility; they exist and are strictly decreasing because U_s are strictly concave increasing.

It is proved in [10] that under the basic algorithm (3) and (4), $\lim_{t \rightarrow \infty} x(t) = x^*$, the unique optimal source rate for the primal problem, provided that the step size $\gamma > 0$ is sufficiently small.

C. REM

A straightforward implementation of the basic algorithm can lead to large backlog and delay. Indeed, the buffer occupancy $b(t)$ at the link at time t evolves according to³

$$b(t+1) = \left[b(t) + \sum_s x_s(t) - c \right]^+.$$

Comparing with (3), we see that backlogs are related to prices by $b(t) = \gamma^{-1}p(t)$. The backlog can be quite large in equilibrium since $\gamma > 0$ is typically small.

To correct this situation, we modify in [2] the price adjustment (3) to

$$p(t+1) = \left[p(t) + \gamma \left(\alpha b(t) + \sum_s x_s(t) - c \right) \right]^+ \quad (5)$$

where $\alpha > 0$ is a small constant. Hence the price is increased when the backlog or the rate is large compared with the capacity, and is reduced otherwise. We will refer to the algorithm defined by (4) and (5) as REM.

III. CONVERGENCE

We now prove that REM matches rate and clears buffer. To avoid triviality, we assume $\sum_s m_s < c < \sum_s M_s$.

Theorem 1: Suppose U_s are strictly concave increasing, twice differentiable, and $\min_s, \{|U''_s(x_s)|, m_s \leq x_s \leq M_s\} > 0$. If $(p(t), x(t))$ is a sequence generated by REM (4) and (5), then $x(t) \rightarrow c$, provided that γ is sufficiently small. Moreover $b(t) \rightarrow 0$.

We will first prove convergence with a single source and then show that multiple sources are equivalent to a single source with an appropriate utility function.

A. Proof: Single Source

Suppose that there is only one source with utility function $U(x)$. Obviously, in this case the optimal source rate is $x^* = c$ and the optimal link price is $p^* = U'(c)$. Setting

$$\gamma = \min_{m \leq x \leq M} \{ |U''(x)| \}$$

we show that $\lim_{t \rightarrow \infty} x(t) = c$ and $\lim_{t \rightarrow \infty} b(t) = 0$.

The key observation is that

$$x(t) \leq c \Rightarrow x(t+1) \leq c. \quad (6)$$

³This is an approximation as it assumes the input rate at link l equals the aggregate source rate

To see this, notice that $x(t) = \max\{U'^{-1}(p(t)), m\} < M$ and $U'^{-1}(\cdot)$ is decreasing. Then $p(t) \geq U'(x(t))$. Therefore,

$$\begin{aligned} U'(c) &= U'(x(t)) + U''(\xi)(c - x(t)) \\ &\leq p(t) + U''(\xi)(c - x(t)) \\ &\leq p(t) + \gamma(x(t) - c) \\ &\leq p(t) + \gamma(\alpha b(t) + x(t) - c) = p(t+1) \end{aligned}$$

where $x(t) < \xi < c$. This implies (6).

If at some time t_0 we have $x(t_0) \leq c$ and $b(t_0) = 0$ then we have $p(t_0 + 1) \leq p(t_0)$ and hence, in addition to (6), we have $x(t) \leq x(t+1) \leq c$ for all $t \geq t_0$. Thus $\lim_{t \rightarrow \infty} x(t)$ exists and it can be shown that it is c .

If at some time t_0 we have $x(t_0) \leq c$ and $b(t_0) > 0$ then

$$b(t_0 + 1) = [b(t_0) + x(t_0) - c]^+ \leq b(t_0).$$

This, together with (6), implies that $b(t)$ is a decreasing sequence, so that $\lim_{t \rightarrow \infty} b(t) = b^*$ exists. Then either $b(t_1) = 0$ and $x(t_1) \leq c$ [because of (6)] at some $t_1 > t_0$, which is then the case discussed above, or $b(t) > 0$ for all $t \geq t_0$. In the latter case, we have $b(t+1) = b(t) + x(t) - c$. Thus

$$\lim_{t \rightarrow \infty} (x(t) - c) = \lim_{t \rightarrow \infty} (b(t+1) - b(t)) = 0.$$

We point out that we must have $b^* = 0$. If $b^* > 0$, since $x(t) \rightarrow c$ when $t \rightarrow \infty$, there exists t_1 , for all $t \geq t_1$, we have $c - x(t) < \alpha b^*/2$. Therefore, for $t \geq t_1$,

$$\begin{aligned} p(t+1) &= p(t) + \gamma(\alpha b(t) + x(t) - c) \\ &\geq p(t) + \gamma \alpha b^*/2 \end{aligned}$$

which yields $p(t) \rightarrow \infty$ when $t \rightarrow \infty$. This contradicts $\lim_{t \rightarrow \infty} x(t) = c$.

Lastly suppose that $x(t_0) \geq c$ and $b(t_0) > 0$ at some t_0 . Then

$$p(t_0 + 1) = p(t_0) + \gamma(b(t_0) + x(t_0) - c).$$

Then at some time $t_1 > t_0$, we will have $x(t_1) < c$, which is then covered by the two cases discussed above. Otherwise, the link price $p(t)$ will approach infinity and $x(t) \geq c$, for all $t \geq t_0$, a contradiction.

We have thus shown that in the single-link single-source case, the input rate under REM approaches the link capacity and the buffer size approaches zero.

B. Proof: Multiple Sources

Assume that there are $S \geq 2$ sources sharing the same link and the utility function of source s is $U_s(x_s)$. We convert the problem into the single-source case.

We define a function $y(p)$ for $p \in [0, \infty)$ as follows. Recall that given price p , source s chooses the rate

$$x_s(p) = [U'_s{}^{-1}(p)]_{m_s}^{M_s}.$$

Define

$$y(p) = \sum_{s=1}^S x_s(p).$$

Then $y(p)$ is nonincreasing. Let

$$u = \min\{U'_s(M_s), 1 \leq s \leq S\}$$

$$v = \max\{U'_s(m_s), 1 \leq s \leq S\}.$$

Since we have $U''_s(x) < 0$, it is easy to see that, for $p \in [u, v]$, $y(p)$ is strictly decreasing and hence the inverse function exists. Use $f(y)$ to denote the inverse function. For y satisfying $\sum_s m_s \leq y \leq \sum_s M_s$, define

$$V(y) = \int_m^y f(t) dt.$$

Then V is strictly increasing, twice differentiable with exception of at most $2S - 2$ points, and $V''(y) < 0$. Obviously, the primal problem is equivalent to maximize $V(y)$ subject to the condition $y \leq c$. This in turn is equivalent to the case of a single source with utility function $V(y)$.

IV. CONCLUDING REMARK

We have not found a precise proof for the multilink case. With a single link, once the rate $x(t) \leq c$, then either $x(t) \uparrow c$ or $b(t) \downarrow 0$. In the multilink case we lose this monotonicity so that the method no longer applies. In [11], an elegant Lyapunov argument is given that proves the *global* convergence of the algo-

rithm, modeled as a continuous time system, in a multilink network. In [1], we present a local convergence proof for a discrete time multilink network. These proofs confirm our empirical experience that REM matches rate and clears buffer, as observed in [2].

REFERENCES

- [1] S. Athuraliya, V. H. Li, S. H. Low, and Q. Yin. REM: Active queue management [Online]. Available: <http://netlab.caltech.edu>
- [2] S. Athuraliya and S. Low. (2000, May) Optimization flow control, II: Random exponential marking. [Online]. Available: <http://netlab.caltech.edu>
- [3] L. S. Brakmo and L. L. Peterson, "TCP Vegas: End to end congestion avoidance on a global Internet," *IEEE J. Select. Areas Commun.*, vol. 13, Oct. 1995.
- [4] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Trans. Networking*, vol. 1, no. 4, pp. 397–413, Aug. 1993.
- [5] V. Jacobson, "Congestion avoidance and control," in *Proc. SIGCOMM'88, ACM*, Aug. 1988, <ftp://ftp.ee.lbl.gov/papers/congavoid.ps.Z>.
- [6] F. P. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.
- [7] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, Mar. 1998.
- [8] S. Kunnipur and R. Srikant, "End-to-end congestion control schemes: Utility functions, random losses and ECN marks," in *Proc. IEEE Infocom*, Mar. 2000, <http://www.ieee-infocom.org/2000/papers/401.ps>.
- [9] S. H. Low, "A duality model of TCP flow controls," in *Proc. ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management*, Sept. 18–20, 2000.
- [10] S. H. Low and D. E. Lapsley, "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, pp. 861–874, Dec. 1999.
- [11] F. Paganini, "On the stability of optimization-based flow control," submitted for publication.
- [12] W. Stevens, *TCP/IP Illustrated: The Protocols*. Reading, MA: Addison-Wesley, 1999, vol. 1, 15th printing.