Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

On nonlinear AIMD congestion control for high-speed networks

Robert Shorten*, Fabian Wirth*, Mehmet Akar*

Abstract— **This papers analyzses a class of nonlinear positive systems that model the dynamics of nonlinear *additive-increase multiplicative-decrease* (AIMD) protocols. The system class covers a range of protocols that are currently used in real communication networks, such as standard TCP, and recent proposals for congestion control protocols such as Scalable TCP.**

I. INTRODUCTION

Traffic generated by the *Transmission Control Protocol* (TCP) accounts for 85% to 95% of all traffic in today's internet [1]. TCP, in congestion avoidance mode, is based primarily on the Chiu and Jain's [2] *Additive-Increase Multiplicative-Decrease* (AIMD) paradigm for decentralized allocation of a shared resource (e.g., bandwidth) among competing users. The AIMD paradigm is based upon a network of users competing for the available resource by using two basic strategies; they probe for their share of the available resource by utilising more and more of the resource (the additive increase stage), and then instantaneously down-scale their utilization-rates in a multiplicative fashion when notified (simultaneously) that capacity was reached (the multiplicative decrease stage). With some minor modifications, the AIMD algorithm has served the networking community well over the past two decades and it continues to provide the basic building block upon which today's internet communication is built.

Recently, in the context of designing high speed communication networks, several authors have suggested basic modifications to the AIMD algorithm; for example, see [3], [4], [5], [6], [7]. One idea underlying these modifications is to allow the employment of more aggressive probing for available bandwidth by replacing the linear in time increase of probing that is a feature of TCP with nonlinear growth. We refer to these algorithms as nonlinear AIMD (NAIMD) algorithms. While the modifications appear minor from an algorithmic viewpoint, they result in networks with different dynamic properties than those employing the basic (linear) AIMD; see [8]. Remarkably, despite increasing deployment of these algorithms many basic questions pertaining to the behavior of such networks have not yet been addressed.

The objective of the current paper is to examine and study a class of NAIMD algorithms. Under the assumption of user-synchronization, necessary and sufficient conditions are derived which guarantee that the network has a unique stable outcome to which it converges geometrically under all starting points.

* Hamilton Institute, NUI Maynooth, Maynooth, Co. Kildare, Ireland {Robert.Shorten, Fabian.Wirth, Mehmet.Akar}@nuim.ie

II. PREAMBLE: AIMD CONGESTION CONTROL

In their original paper [2], Chiu and Jain consider a system in which n users compete for a resource having limited availability per unit time, e.g., bandwidth in communication networks. The users' actions consist of (continuously) probing the availability of the resource by submitting requests for its use – these requests are satisfied whenever global capacity is not exceeded. A key assumption in the model formulated by Chiu and Jain is the assertion that the users do not communicate directly with each other. Further, the only information about availability of the resource that the users get is when the collective utilization of the resource exceeds some capacity constraint. At such time-instances, referred to as *congestion events*, all users are instantly and simultaneously informed through a binary feedback. The users are assumed to respond instantly to these notifications of congestion by decentralized down-scaling of their individual utilization-rates. Given this basic setting, the problem is then to develop an algorithm that produces probing strategies for the users so that each user will infer its "fair" share of the shared resource in a decentralized manner.

In the current paper we focus on the *synchronized* problem, referring to simultaneous notification of congestion to all users to which they all respond. In *unsynchronized systems*, the signal about system-saturation is not transmitted simultaneously to all users. While synchronization is not valid in many real communication networks, the study of such systems is important for two reasons. First, it represents an important first step towards the understanding of more general systems. And Second, synchronization appears to be a common feature of high speed communication networks [7] and consequently the understanding the behavior of such networks may be of merit in some practical situations.

III. LINEAR AIMD CONGESTION CONTROL

The AIMD algorithm of Chiu and Jain describes probing strategies that evolve in cycles, each cycle having two phases. The first phase of the cycles is instantaneous. It occurs when capacity is reached, users are notified and each responds by down-scaling its utilization-rate (abruptly) by a multiplicative factor. This phase is called the *Multiplicative Decrease* (MD) phase. During the second phase of a cycle, each user increases his utilization-rate linearly until congestion is reached again, at which point the first phase of the next cycle is entered. The second phase is called the *Additive Increase* (AI) phase.

Denote the share of the collective resource allocated to user i at time t by $x_i(t)$ and let $x(t) = [x_1(t), \ldots, x_n(t)]^T$. The capacity constraint requires that $\sum_{i=1}^{n} x_i(t) < C$, with C as the total capacity of the resource available to the entire system.

The k^{th} cycle begins at a time $t(k)$ at which the global utilization of the resource reaches capacity. The instantaneous decrease of the utilization-rate of user i during the MD phase of the k^{th} cycle is expressed by:

$$x_i(t(k)^+) = \beta_i x_i(t(k)^-), \quad (1)$$

where $x_i(t(k)^+) := \lim_{t \searrow t(k)} x_i(t)$, $x_i(t(k)^-) := \lim_{t \nearrow t(k)} x_i(t)$ and β_i is a constant in the open interval $(0, 1)$. During the AI phase of the k^{th} cycle, the utilization-rate of user i evolves according to:

$$x_i(t) = x_i(t(k)^+) + \alpha_i(t - t(k)), \quad (2)$$

where α_i is a positive constant. The $(k+1)^{st}$ cycle begins at time $t(k+1)^+$ that equals the time t for which the right hand-side of (2) reaches capacity.

A convenient framework to study the implication of the AIMD algorithm is to consider the utilization-rates at congestion events that occur at times $t^1, t^2, \ldots$. Combining (1) and (2), we see that the evolution of the utilization-rate of user i between the k^{th} and $(k+1)^{st}$ congestion points is given by

$$x_i(t(k+1)) = \beta_i x_i(t(k)) + \alpha_i\Big(t(k+1) - t(k)\Big), \quad (3)$$

This avenue of investigation is explored in [9] where it is shown that the transformation of the utilization-rates between consecutive congestion points is linear with the representation

$$x(k+1) = Ax(k) \quad (4)$$

where

$$A = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} + \frac{1}{\sum_{j=1}^{n} \alpha_i} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} 1-\beta_1, & \cdots & ,1-\beta_n \end{bmatrix}. \quad (5)$$

IV. Nonlinear AIMD Congestion Control

We next describe a nonlinear variant of the basic AIMD algorithm, to which we refer as NAIMD. Specifically, NAIMD coincides with the standard (linear) AIMD except that in the AI phase the increase in the utilization-rate of each user i is dictated by a nonlinear function of the state $x_i(k)$ that we denote $a_i(x_i(k))$. In this paper we assume that $a_i(x_i(k)) = x_i(k)^{\kappa_i}(t(k+1) - t(k))$ where we assume that $x_i(k) > 0$ and where $\kappa_i \in \mathbb{R}$. This functional form is motivated by the desire to make TCP more aggressive in high-speed and long distance networks. The evolution of the utilization-rate of user i between the k^{th} and $(k+1)^{st}$ congestion points is then given by

$$x_i(k+1) = \beta_i x_i(k) + (x_i(k))^{\kappa_i}(t(k+1) - t(k)), \quad (6)$$

which replaces (3). This family of congestion control protocols includes standard TCP when $\kappa = 0$, Scalable TCP when $\kappa = 1$, and several other proposed algorithms for high speed networks.

A slight and natural generalization is to use the function $a_i(x_i(k)) = c_i x_i(k)^{\kappa_i}(t(k+1)-t(k))$ for some constants $c_i > 0$. The results in the present paper can be easily generalized to this more general situation. Namely, by introducing the constants $d_i = c_i^{1-\alpha_i}$ and the transformed variables $z_i = d_i x_i$, $i = 1, \ldots, n$ it is easy to see that the dynamic equation of z are of the form (6). However, the constraint condition changes to $C = \sum d_i^{-1} z_i$, so that in full generality we have to study (6) with a general linear constraint. The effect on the results however is marginal and we therefore do not pursue this issue here.

It has been recently shown by several authors that some choices of the $a_i(t)$ lead to poor dynamic properties, including the lack of stable utilization-rates; see [8]. It is therefore of interest to determine the properties of networks with various values of κ.

The next example illustrates the evolution of the utilization rates in a 2-user system in which NAIMD is applied and to which our forthcoming results apply.

Example 1: Consider a system with 2-users that apply NAIMD with $\beta_1 = 0.25, \beta_2 = 0.5$ and $\kappa_1 = \kappa_2 = 1$. This is Scalable TCP. The resulting network behaviour is unstable as one user takes all of the resource.

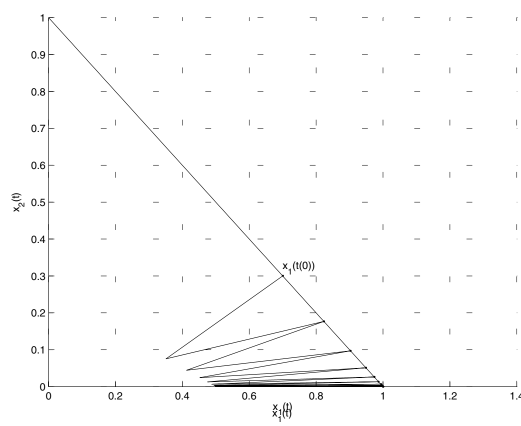

Fig. 1. Utilization-rates under NAIMD with $\kappa = 1$.

V. A positive systems model of NAIMD algorithms

Throughout, the following notation is adopted: $\mathbb{R}$ denotes the real numbers; $\mathbb{R}^n$ denotes the n-dimensional real Euclidean space; $\mathbb{R}^{n \times n}$ denotes the space of $n \times n$ matrices with real entries; x_i denotes the i^{th} component of the vector x in $\mathbb{R}^n$. We denote the i-th unit vector by e_i. The positive orthant is denoted by $\mathbb{R}_+^n := \{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \ldots, n\}$ and we write $x > 0$ if $x_i > 0, i = 1, \ldots, n$. The absolute value of a vector $x \in \mathbb{R}^n$ is defined by $|x| = \begin{bmatrix} |x_1| & \ldots & |x_n| \end{bmatrix}^T$.

We consider a network of n users each with a state $x_i(k)$, competing for a shared resource. We assume that each user behaves according to the following strategy. For each $i \in \{1, ..., n\}$ source i adapts its window size according to

$$x_i(k+1) = \beta_i x_i(k) + x_i(k)^{\kappa_i}(t(k+1) - t(k)), \quad (7)$$

subject to the constraint that $\sum_{i=1}^{n} x_i(k) = C$ for all k, where $\beta_i \in (0, 1)$ and $C > 0$ denotes the capacity of the resource. The index k enumerates the time instances at which the capacity constraint is reached and $\beta_i \in (0,1)$ is the backoff parameter applied when the i'th source is informed of 'congestion'. The constants $\kappa_i, i = 1, \ldots, n$ determine the aggressiveness of each of the sources.

The time T between congestion event depends on the state at the beginning of the event and is given by $(t(k+1) - t(k)) = T(x(k))$, where

$$T(x) = \frac{C - \sum_{j=1}^{n} \beta_j x_j}{\sum_{j=1}^{n} x_j^{\kappa_j}} = \frac{\sum_{j=1}^{n}(1-\beta_j)x_j}{\sum_{j=1}^{n} x_j^{\kappa_j}}. \tag{8}$$

To write the overall map describing the system compactly we introduce the diagonal matrix $D_\beta := \mathrm{diag}(\beta_1, \ldots, \beta_n) \in \mathbb{R}^{n\times n}$ and we write $x^{\cdot\kappa} \in \mathbb{R}^n$ defined by $x = (x_1, \ldots, x_n)^T \mapsto (x_1^{\kappa_1}, \ldots, x_n^{\kappa_n})^T =: x^{\cdot\kappa}$. With this we obtain

$$\begin{aligned} x(k+1) &= f(x(k)), \text{ where} \\ f(x) &= D_\beta x + T(x) x^{\cdot\kappa}. \end{aligned} \tag{9}$$

All n equations can be written compactly in matrix form

$$x(k+1) = A(x(k))x(k), \tag{10}$$

where

$$\begin{aligned} A(x) &= \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} \\ &+ \frac{1}{\sum_{j=1}^{n} x_j^{\kappa_j}} \begin{bmatrix} x_1^{\kappa_1} \\ x_2^{\kappa_2} \\ \cdots \\ x_n^{\kappa_n} \end{bmatrix} \begin{bmatrix} 1-\beta_1 \cdots 1-\beta_n \end{bmatrix}. \end{aligned} \tag{11}$$

It is readily verified that for each vector of network states x the matrix $A(x)$ is column stochastic, i.e it is a matrix with nonnegative entries, whose columns sum to 1. We now introduce the stability property of interest for networks of the type that has been described.

We note that, unless $\kappa_i = 0$, the model prevents source i from ever increasing its share, if the initial condition is $x_i(0) = 0$. We will therefore always assume that the initial condition satisfies $x(0) > 0$ and we are interested in the dynamics on the invariant set

$$\tilde{\Sigma} := \{x \in \mathbb{R}^n \mid x > 0, \sum x_i = C\}.$$

In practice, in AIMD algorithms there is usually a special "slow start" procedure which ensures that a source is sending at a positive rate, before the actual AIMD algorithm is used. We aim to analyse the dynamics of NAIMD and assume therefore, that the system has already performed slow start.

We now present the definition of stability that is investigated in this paper. From the perspective of network design it is desirable, that the system has a unique fixed to which all initial conditions converge. This is th egist of the following

Definition 5.1 (Stable Networks): A network given by dynamics of the form (7) on $\tilde{\Sigma}$ is called stable if

(i) system (7) has a unique equilibrium point $w^* \in \tilde{\Sigma}$,
(ii) w^* is a globally asymptotically stable fixed point of (7) on $\tilde{\Sigma}$.

A network is said to be unstable if it is not stable. □

Remark 5.1: Thus by definition a network is unstable if at least one user is starved of the resource asymptotically or if at least one user does not asymptotically achieve a constant share of the resource.

VI. Fixed points, linearization, stability, and instability results

In this section we perform the first step in our analysis and identify the parameters for which there is a fixed point and analyse the local stability properties of this fiped point with linearization techniques.

A. Fixed Points

Let us first determine the fixed points of (9). Recall, that we are under the assumption that if congestion occurs then $\sum_j x_j = C > 0$.

Proposition 6.1: Consider the system (9), (8). Then

(i) if $(1-\kappa_i)(1-\kappa_j) > 0$, for all $i, j \in \{1, \ldots, n\}$ then there exists a unique fixed point $w^* \in \tilde{\Sigma}$. This fixed point is determined by the equations

$$w_i^* = \left(\frac{T(w^*)}{(1-\beta_i)}\right)^{1/(1-\kappa_i)}, \quad i = 1, \ldots, n, \tag{12}$$

where $T(w^*)$ is the unique solution of

$$C = \sum_{j=1}^{n} \left(\frac{T^*}{(1-\beta_j)}\right)^{1/(1-\kappa_j)}. \tag{13}$$

(ii) if $\kappa_i \neq 1, i = 1, \ldots, n$ and there are $i, j \in \{1, \ldots, n\}$, such that $(1-\kappa_i)(1-\kappa_j) < 0$ then the system is generically unstable, that is, there exists a constant $C^* > 0$, such that the system (9), (8) has no fixed point in $\tilde{\Sigma}$, if $0 < C < C^*$ and multiple fixed points, if $C > C^*$.

(iii) if exactly one source i satisfies $\kappa_i = 1$, then there exists a unique fixed point $w^* \in \tilde{\Sigma}$ if and only if

$$C - \sum_{j\neq i} \left(\frac{1-\beta_i}{1-\beta_j}\right)^{1/(1-\kappa_j)} > 0. \tag{14}$$

(iv) if two or more sources sources $i_1, \ldots, i_l$ have the parameter $\kappa_{i_1} = \ldots = \kappa_{i_l} = 1$, then the network is unstable. In particular,
  (a) if $\beta_{i_\nu} \neq \beta_{i_\mu}$ for some $\nu, \mu \in \{1, \ldots, l\}$ then the system has no fixed points in $\tilde{\Sigma}$,
  (b) if $\beta_{i_1} = \ldots = \beta_{i_l}$ then the system has no or more than one fixed point in $\tilde{\Sigma}$.

Remark 6.1: Note that the instability behaviour observed in Example 1 is explained by the statement of Proposition 6.1 (iv) (a).

Proof: (i) Assume $(1-\kappa_i)(1-\kappa_j) > 0, i, j = 1, \ldots, n$. From the fixed point equations

$$w_i^* = \beta_i w_i^* + T(w^*) w_i^{*\kappa_i}, \quad i = 1, \ldots, n \tag{15}$$

we obtain the condition

$$w_i^* = \left(\frac{T(w^*)}{(1-\beta_i)}\right)^{1/(1-\kappa_i)}, \quad i = 1, \ldots, n. \tag{16}$$

In addition to (16) the fixed point has to satisfy the condition $\sum_{j=1}^n w_j^* = C$. Thus the constant $T^* := T(w^*)$ has to satisfy

$$C = \sum_{j=1}^n w_j^* = \sum_{j=1}^n \left(\frac{T^*}{(1-\beta_j)}\right)^{1/(1-\kappa_j)} . \qquad (17)$$

Now by assumption the powers $1/(1-\kappa_j), j = 1, \dots, n$ are either all negative or all positive. In both cases, the map

$$T \mapsto \sum_{j=1}^n \left(\frac{T}{(1-\beta_j)}\right)^{1/(1-\kappa_j)}$$

is a homeomorphism of $(0, \infty)$ so that for all $C > 0$ there is a unique T^* satisfying (17). This together with (16) determines a unique fixed point in $\tilde{\Sigma}$.
(ii) As in (i) we obtain the conditions (16) and (17) for a fixed point. Consider again the map

$$T \mapsto \sum_{j=1}^n \left(\frac{T}{(1-\beta_j)}\right)^{1/(1-\kappa_j)} .$$

As there are now indeces i, j such that $1/(1-\kappa_i), 1/(1-\kappa_j)$ have a different sign, this means that this map tends to infinity for $T \to 0$ and $T \to \infty$. Denoting

$$C^* := \min_{T>0} \sum_{j=1}^n \left(\frac{T}{(1-\beta_j)}\right)^{1/(1-\kappa_j)}$$

this shows that for $0 < C < C^*$ there is no solution of (16) and (17), whereas for $C > C^*$ there is more than one solution of these conditions.
(iii) We now assume that $\kappa_1 = 1$, $\kappa_i \neq 1$, $i = 2, \dots, n$. Assuming that there is a fixed point $w^* \in \tilde{\Sigma}$ the fixed point equation (15) for $i = 1$ reduces to

$$T(w^*) = (1-\beta_1) . \qquad (18)$$

Thus we obtain from (16) that for $i = 2, \dots, n$ we have

$$w_i^* = \left(\frac{1-\beta_1}{1-\beta_i}\right)^{1/(1-\kappa_i)} .$$

Now $w^* \in \tilde{\Sigma}$ in conjunction with $\sum_{j=1}^n w_j^* = C$ is possible, if and only if (14) holds, as desired.
(iv) Assume now, that $\kappa_1 = \dots = \kappa_l = 1$, $\kappa_i \neq 1$, $i = l+1, \dots, n$. Arguing just as before, we obtain from (18), that if a fixed point $w^* \in \tilde{\Sigma}$ exists, then $T(w^*) = (1-\beta_1) = \dots = (1-\beta_l)$. This shows part (a) of the claim. Assuming that the β_i, $i = 1, \dots, l$ coincide we see that a necessary condition for the existence of a fixed point in $\tilde{\Sigma}$ is that

$$d := C - \sum_{j=l+1}^n \left(\frac{1-\beta_1}{1-\beta_j}\right)^{1/(1-\kappa_j)} > 0 .$$

If this is not the case no fixed point in $\tilde{\Sigma}$ exists. Otherwise any choice of $w_1, \dots, w_l > 0$ such that $\sum_{j=1}^l w_j = d$ leads to a fixed point, so that the fixed points in $\tilde{\Sigma}$ are not unique. ■

We note for further reference that from (16) it follows for all κ_i, that if a fixed point w^* in $\tilde{\Sigma}$ exists then

$$(1-\beta_i) = T(w^*) w_i^{*\kappa_i - 1} , \quad i = 1, \dots, n . \qquad (19)$$

B. Linearization and local stability results

In the analysis of the existence of fixed points we have seen, that there is essentially only one situation of interest, which is the case that $(1-\kappa_i)(1-\kappa_j) > 0$ for all $i, j = 1, \dots, n$. In all other cases, either the system cannot be stable, as there are no or many fixed points in $\tilde{\Sigma}$, or sometimes by a particular choice of C, or by allowing exactly one user to use the parameter $\kappa_i = 1$ a unique fixed point is obtained. As the latter cases are not particularly relevant (how would one pick that one user for instance?) we do not discuss these any further.
We now investigate local stability of the unique fixed point, that exists, if the condition $(1-\kappa_i)(1-\kappa_j) > 0$ for all $i, j = 1, \dots, n$ is satisfied. In the following statement we abbreviate

$$p := \frac{w^{*\cdot\kappa}}{\sum_{j=1}^n w_j^{*\kappa_j}} \in \mathbb{R}^n ,$$

and we introduce

$$\begin{aligned} \gamma_i &:= \beta_i + (1-\beta_i)\kappa_i , \quad \gamma := \begin{bmatrix}\gamma_1 & \dots & \gamma_n\end{bmatrix}^T , (20) \\ D_\gamma &= \operatorname{diag}(\gamma_1, \dots, \gamma_n) \in \mathbb{R}^{n\times n} . \end{aligned}$$

In the sequel we assume without loss of generality, that the γ_i are ordered

$$\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_n .$$

Note that $p > 0$ is a stochastic vector, i.e. its entries sum to 1.
Proposition 6.2: Consider the system (9), (8) and let $(1-\kappa_i)(1-\kappa_j) > 0$, for all $i, j \in \{1, \dots, n\}$. Then the Jacobian Jf of f in the unique fixed point w^* of (9) is given by

$$Jf(w^*) = D_\gamma - p\gamma^T . \qquad (21)$$

Proof: Considering (9), calculation of the Jacobian of f in x leads to

$$\begin{aligned} Jf(x) &= D_\beta + T(x)\operatorname{diag}(\kappa_1 x_1^{\kappa_1 - 1}, \dots, \kappa_n x_n^{\kappa_n - 1}) \\ &\quad + x^{\cdot\kappa}\nabla T(x) \\ &= D_\beta + T(x)\operatorname{diag}(\kappa_1 x_1^{\kappa_1 - 1}, \dots, \kappa_n x_n^{\kappa_n - 1}) \\ &\quad + \frac{x^{\cdot\kappa}}{\sum_{j=1}^n x_j^{\kappa_j}} \left(\begin{bmatrix}-\beta_1 & \dots & -\beta_n\end{bmatrix}\right. \\ &\quad \left. - T(x)\begin{bmatrix}\kappa_1 x_1^{\kappa_1-1} & \dots & \kappa_n x_n^{\kappa_n-1}\end{bmatrix}\right) , \end{aligned}$$

where we have used (8). Inserting the fixed point w^* into this equation we obtain (21) using (19) and (20). ■
We note that the column sums of $Jf(w^*)$ are all equal to 0, thus 0 is an eigenvalue of $Jf(w^*)$ with associated left, resp. right eigenvectors

$$\begin{bmatrix}1 & \dots & 1\end{bmatrix} ,$$

and (if $\gamma_i \neq 0, i = 1, \dots, n$)

$$\begin{bmatrix}\frac{p_1}{\gamma_1} & \dots & \frac{p_n}{\gamma_n}\end{bmatrix}^T .$$

We will now investigate the location of the eigenvalues of $Jf(w^*)$.
Theorem 6.1: Consider the system (9), (8) and let $(1-\kappa_i)(1-\kappa_j) > 0$, for all $i, j \in \{1, \dots, n\}$. The following statements hold:

(i) The spectrum of $Jf(w^*)$ is real.

(ii) The spectral radius $r(Jf(w^*))$ satisfies

$$r(Jf(w^*)) \in (\gamma_{n-1}, \gamma_n)\,, \quad \text{if } \gamma_n \geq |\gamma_1|\,,$$
$$r(Jf(w^*)) \in (|\gamma_2|, |\gamma_1|)\,, \quad \text{if } \gamma_n \leq |\gamma_1|\,.$$

Proof: We prove the result under the assumption, that $\gamma_i \neq 0$, $i = 1, \ldots, n$. As the entries of $Jf(w^*)$ depend continuously on γ it then follows using perturbation results, that the statements also hold for the case, that some γ_i vanish.
(i) We first note that if $\eta = \gamma_j = \ldots = \gamma_{j+l}$ for some $j = 1, \ldots, n$, $l > 0$ then η is an eigenvalue of multiplicity l of $Jf(w^*)$. To see this consider constants $a_j, \ldots, a_{j+l}$ not all zero, such that $\sum_{m=0}^{l} a_{j+m} = 0$. Then for the nonzero vector $y := \sum_{m=0}^{l} a_{j+m} e_{j+m}$ we have

$$Jf(w^*)y = D_\gamma y - p\gamma^T y = \eta y + p(\eta \sum_{m=0}^{l} a_{j+m}) = \eta y\,,$$

where we have used, that $\gamma_j = \eta$ for all indices j for which $y_j \neq 0$. The dimension of the space of vectors y which can be constructed in this manner is l, showing the first assertion.
Consider now the eigenvalue equation

$$(D_\gamma - p\gamma^T)x = \lambda x$$

If $\gamma^T x = 0$, this implies that x is an eigenvector of D_γ, so that $\lambda = \gamma_i$ for some i. Assume now that $\gamma^T x \neq 0$, so that we may assume it is equal to 1. Then the eigenvalue equation reads in componentwise form

$$\gamma_i x_i - p_i = \lambda x_i\,, \quad i = 1, \ldots, n\,. \tag{22}$$

From this we see that $\lambda = \gamma_i$ for some i if and only if $p_i = 0, i = 1, \ldots, n$, which contradicts $p > 0$. Thus the assumption $\gamma^T x \neq 0$ implies, that λ is different from the γ_i. Thus (22) is equivalent to

$$x_i = \frac{p_i}{\gamma_i - \lambda}\,, \quad i = 1, \ldots, n\,.$$

As we have the condition $\gamma^T x = 1$ we obtain

$$1 = \sum \gamma_i x_i = \sum \frac{\gamma_i p_i}{\gamma_i - \lambda} =: q(\lambda)\,.$$

By these consideration we see, that λ is an eigenvalue different from the γ_i, if and only if $q(\lambda) = 1$.
Clearly, the rational function q has poles in $\gamma_1, \ldots, \gamma_n$. Note that as η approaches γ_i from above we have $q(\eta) \to -\text{sign}\,(\gamma_i)\infty$ and for η approaching γ_i from below, it holds that $q(\eta) \to \text{sign}\,(\gamma_i)\infty$. Thus if $\text{sign}\,(\gamma_i) = \text{sign}\,(\gamma_{i+1})$ we have that q maps the interval (γ_i, γ_{i+1}) to $\mathbb{R}$ and consequently, there is a $\gamma_i < \lambda < \gamma_{i+1}$ with $q(\lambda) = 1$, so that λ is an eigenvalue of $Jf(w^*)$. If $\text{sign}\,(\gamma_i) \neq \text{sign}\,(\gamma_{i+1})$, that is, if $\gamma_i < 0 < \gamma_{i+1}$, then the eigenvalue 0 is in the interval (γ_i, γ_{i+1}). Thus in total every interval (γ_i, γ_{i+1}) contains an eigenvalue, if we set $(\gamma_i, \gamma_{i+1}) = \{\gamma_{i+1}\}$ for the degenerate case $\gamma_i = \gamma_{i+1}$.
In all we have seen up to now, that $n-1$ of the eigenvalues of $Jf(w^*)$ are real, which implies that all eigenvalues of $Jf(w^*)$ are real, because the matrix is real.
(ii) To complete the proof, note that the $n-1$ eigenvalues we have accounted for lie in the interval $[\gamma_1, \gamma_n]$.
If $\gamma_1 > 0$, then the eigenvalue 0 is an additional eigenvalue, so that all eigenvalues lie in the interval $[0, \gamma_n]$ and one eigenvalue is in the interval $[\gamma_{n-1}, \gamma_n]$. This eigenvalue is equal to the spectral radius $r(Jf(w^*))$, which shows the claim in this case. Similarly, if $\gamma_n < 0$, then 0 is the remaining eigenvalue and the spectral radius satisfies $r(Jf(w^*)) \in [\,|\gamma_2|, |\gamma_1|\,]$. Finally, if $\gamma_1 < 0 < \gamma_n$, then there is an index i, such that $\gamma_i < 0 < \gamma_{i+1}$. We will show that in this case there are two eigenvalues in the interval (γ_i, γ_{i+1}) which again shows the claim concerning the spectral radius. In this case either $|\gamma_1|$ or γ_n could be the upper bound for the spectral radius depending on their relative size.
To show, that there are two eigenvalues in the interval (γ_i, γ_{i+1}), note that for η lower than γ_i we have $q(\eta) \to \infty$, because $\text{sign}\,(\gamma_i) = -1$ and for η greater than γ_{i+1} we have $q(\eta) \to \infty$. Note that $q(0) = 1$, so that $-\infty < \min_{\eta \in (\gamma_i, \gamma_{i+1})} q(\eta) \leq 1$. If $\min_{\eta \in (\gamma_i, \gamma_{i+1})} q(\eta) < 1$, then necessarily $q(\lambda_i) = 1$ for two different values λ_1, λ_2, as q approaches ∞ at both ends of the interval. If $\min_{\eta \in (\gamma_i, \gamma_{i+1})} q(\eta) = 1$, then we can again argue by perturbation analysis: As $\lambda = 0$ is a minimum, we have that

$$0 = \frac{d}{d\eta} q(\eta)_{\eta=0} = -\sum_{j=1}^{n} \frac{p_j}{\gamma_j}\,. \tag{23}$$

By considering $\tilde{\gamma}_i := \gamma_i + \varepsilon$, we see, that (23) does not hold for $\tilde{\gamma} := \gamma + \varepsilon\, e_i$ for all $\epsilon > 0$ small enough (because $p > 0$). Applying the previous considerations, the matrix $D_{\tilde{\gamma}} - p\tilde{\gamma}$ has two distinct eigenvalues in the interval $(\tilde{\gamma}_i, \gamma_{i+1})$. Letting $\varepsilon \to 0$ it follows that $Jf(w^*)$ has two eigenvalues in the interval (γ_i, γ_{i+1}), as desired. ■
From these inequalities we obtain the following stability and instability results. Again note, that we only need to concern ourselves with the case $(1-\kappa_i)(1-\kappa_j) > 0$ for all $i, j = 1, \ldots, n$, in which case we may speak of the unique fixed point of the system.
Corollary 6.1: Consider the system (9), (8) and let $\beta_i \in [0, 1), i = 1, \ldots, n$ and $C > 0$.

(i) If $\kappa_i > 1, i = 1, \ldots, n$, then the fixed point $w^* = (w_1^*, \ldots, w_n^*)$ given by (12) is unstable. In particular, the network is unstable.
(ii) If $-(1+\beta_i)/(1-\beta_i) \leq \kappa_i < 1$, $i = 1, \ldots, n$, then the fixed point is locally asymptotically stable.
(iii) If $\kappa_i < 1, i = 1, \ldots, n$, and $\kappa_j < -(1+\beta_i)/(1-\beta_i)$ for two values of $j \in \{1, \ldots, n\}$ then the fixed point is unstable. In particular, the network is unstable.

Proof: (i) This is an immediate consequence of Theorem 6.1 (ii), as $r(Jf(w^*)) > \gamma_2 = \beta_2 + (1-\beta_2)\kappa_2 > 1$, which follows using $\kappa_2 > 1$.
(ii) Note that $-(1+\beta_i)/(1-\beta_i) \leq \kappa_i < 1$ if and only if $\gamma_i = \beta_i + (1-\beta_i)\kappa_i \in (-1, 1)$. This implies by Theorem 6.1 (ii), that $r(Jf(w^*)) \leq \max\{|\gamma_1|, \gamma_n\} < 1$, so that $r(Jf(w^*))$ is Schur stable.
(iii) By assumption $\gamma_2 < -1$ and thus using Theorem 6.1 (ii) it follows that $r(Jf(w^*)) > 1$. Thus the fixed point is unstable. ■

Remark 6.2: (i) Summarizing the only set ups, in which a unique fixed point is achieved, that is locally asymptotically

stable, are given by the choices $-(1+\beta_i)/(1-\beta_i) \le \kappa_i < 1$, $i = 1, \dots, n$.
(ii) Note that the bound $r(Jf(w^*)) \le \max\{|\gamma_1|, \gamma_n\}$ also provides easily computable bounds for the local attraction rates in the case, the fixed point is locally stable.
(ii) We note that in [2, Section 4] also nonlinear protocols of the form (9) are briefly studied. A remark in that section suggests in a heuristic manner that stability of the network will hold for all negative values of the exponent. We see from (iii) in the previous Proposition, that this statement does not hold up to careful scrutiny. In fairness, it has to be admitted, that this remark is only a minor comment in the paper [2], and the authors do not claim to put forward a general result concerning this question.

VII. Global stability

We have so far obtained conditions for the existence of a unique fixed point of the network in terms of the κ_i, namely all κ_i have to be chosen such that the sign of $1 - \kappa_i$ is fixed, (ignoring a few special cases, that are of no particular relevance to the general design of protocols). We have also obtained conditions on the κ_i that guarantee that unique fixed points are locally exponentially stable using linearization theory. To speak of a stable network, however, we would like to achieve global stability. This is the topic of this section.
We now show that this fixed point is asymptotically stable if the constants κ_i all coincide.
Proposition 7.1: Let $\kappa \in [0,1)$, $\kappa_i = \kappa$, $\beta_i \in [0,1), i = 1, \dots, n$ and $C > 0$. Then the fixed point $w^* = (w_1^*, \dots, w_n^*)$ given by (12) is asymptotically stable.

Proof: Let $c := w^*$. Instead of the evolution of $x(k)$ we study the system in the new variable z defined by $x_i = c_i z_i$. This leads to the equations

$$z_i(k+1) = \beta_i z_i(k) + \frac{c_i^{\kappa-1} z_i(k)^\kappa}{\sum_{j=1}^n c_j^\kappa z_j(k)^\kappa} \sum_{j=1}^n (1-\beta_j) c_j z_j(k) \, . \tag{24}$$

Note that by definition the fixed point of (24) is $z^* := \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^T$. Let $M(k) = \max\{z_i(k) \mid i = 1, \dots, n\}$ and $m(k) = \min\{z_i(k) \mid i = 1, \dots, n\}$. We will show that if $m(k) < M(k)$ then we have for all $i = 1, \dots, n$ that

$$\begin{aligned}(1-\beta_i) m(k) < \frac{c_i^{\kappa-1} z_i(k)^\kappa}{\sum_{j=1}^n c_j^\kappa z_j(k)^\kappa} \sum_{j=1}^n (1-\beta_j) c_j z_j(k) \\ < (1-\beta_i) M(k) \, .\end{aligned} \tag{25}$$

Together with (24) this shows for all $i = 1, \dots, n$ that

$$z_i(k+1) \in (m(k), M(k)) \, , \tag{26}$$

or in other words $m(k) < m(k+1) < M(k+1) < M(k)$. It then follows using standard arguments from the theory of Lyapunov functions that $M(k) - m(k) \to 0$, or equivalently $z(k)^T \to \begin{bmatrix} 1 & \dots 1 \end{bmatrix}^T$. This however, is equivalent to $x(k) \to w^*$.
Thus it remains to show that (25) holds. In the following we suppress the dependence of k for the sake of succinctness. Using (19) the left hand side of (25) is equivalent to

$$T^* c_i^{\kappa-1} m < \frac{c_i^{\kappa-1} z_i^\kappa}{\sum_{j=1}^n c_j^\kappa z_j^\kappa} \sum_{j=1}^n T^* c_j^{\kappa-1} c_j z_j \, ,$$

which is equivalent to

$$\sum_{j=1}^n c_j^\kappa \frac{z_j^\kappa}{z_i^\kappa} < \sum_{j=1}^n c_j^\kappa \frac{z_j}{m} \, . \tag{27}$$

Thus the proof of the lower inequality is complete, if we can show that $z_j^\kappa / z_i^\kappa \le z_j/m$ for $j = 1, \dots, n$ with strict inequality for at least one j. We note that $z_i \ge m$ so that if $z_j \ge z_i$, then we have $z_j^\kappa / z_i^\kappa \le z_j/z_i \le z_j/m$. Strict inequality occurs whenever $z_j > z_i$ or when $z_i > m$. One of these has to hold (for an appropriate j), as otherwise $M = m$. On the other hand $z_j/m \ge 1$, so that if $z_j \le z_i$, then $z_j^\kappa / z_i^\kappa \le 1 \le z_j/m$. Again it is easy to see that strict inequality has to hold for some j, if $m < M$. Hence (27) holds. The proof for the right hand equation in (25) follows exactly the same lines. ■

VIII. Conclusions

In this paper we have studied stability properties of heterogeneous networks in which different users implement different versions of nonlinear AIMD algorithms with different levels of aggressiveness. It has been shown that if the level of aggressiveness κ_i is bigger or equal than 1, then this results in an unstable situation. This applies in particular to Scalable TCP, in which users set their aggressiveness to $\kappa = 1$.

Acknowledgements

This work was supported by Science Foundation Ireland grant 04-IN3 -I460.

References

[1] Z. Zhao, S. Darbha, and A. L. N. Reddy, “A method for estimating the proportion of nonresponsive traffic at a router,” *IEEE ACM Trans. Networking*, vol. 12, no. 4, pp. 708–718, 2004.
[2] D. Chiu and R. Jain, “Analysis of the increase and decrease algorithms for congestion avoidance in computer networks,” *Computer Networks and ISDN Systems*, vol. 17, no. 1, pp. 1–14, 1989.
[3] S. Floyd, “High speed TCP for large congestion windows,” tech. rep., Internet draft draft-floyd-tcp-highspeed-02.txt, work in progres, February 2003.
[4] S. Lin, L. Guo, I. Matta, and A. Bestavros, “A spectrum of TCP-friendly window-based congestion control algorithms,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 3, pp. 341–355, 2003.
[5] T. Kelly, “On engineering a stable and scalable TCP variant,” tech. rep., Cambridge University Engineering Department Technical Report CUED/F-INFENG/TR.435, 2002.
[6] D. Leith and R. Shorten, “H-TCP: TCP for high-speed long-distance networks,” in *Proc. 2nd Workshop on Protocols for Fast Long Distance Networks*, (Argonne, IL, USA), February 2004. http://dsd.lbl.gov/DIDC/PFLDnet2004/papers/Leith.pdf.
[7] L. Xu, K. Harfoush, and I. Rhee, “Binary increase congestion control (BIC) for fast long-distance networks,” in *Proc. of IEEE INFOCOM 2004*, (Hong Kong), pp. 2514–2524, 2004.
[8] E. Altman, K. E. Avrachenkov, and B. J. Prabhu, “Fairness in mimd congestion control algorithms,” in *Proc. of IEEE INFOCOM 2005*, (Miami, FL), pp. 1350–1361, 2005.
[9] R. Shorten, F. Wirth, and D. Leith, “A positive systems model of TCP-like congestion control: Asymptotic analysis.” IEEE/ACM Transactions on Networking, in press, 2006.