Internet Congestion Control: Equilibrium and Dynamics

A. Kevin Tang

Cornell University

ISS Seminar, Princeton University, February 21, 2008
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
- Cellular phone networks (Communication/Information Theory)
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
- Cellular phone networks (Communication/Information Theory)
- The Internet
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
- Cellular phone networks (Communication/Information Theory)
- The Internet (???)
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
- Cellular phone networks (Communication/Information Theory)
- The Internet (? ? ?)
  - aggregation of a large number of networks owned by competing entities
  - wide array of technologies and equipment
  - evolve asynchronously and anarchically
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
- Cellular phone networks (Communication/Information Theory)
- The Internet
  - aggregation of a large number of networks owned by competing entities
  - wide array of technologies and equipment
  - evolve asynchronously and anarchically
- A large distributed system that lacks central control!
Networks and Corresponding Theories

- Power networks (Maxwell Theory)
- Telephone networks (Queueing Theory)
- Cellular phone networks (Communication/Information Theory)
- The Internet (aggregation of a large number of networks owned by competing entities; wide array of technologies and equipment; evolve asynchronously and anarchically)
- A large distributed system that lacks central control!
- What global behaviors emerge from the interaction of local algorithms?
Outline and Highlight

• Outline

  Introduction to congestion control and its current theory
  Equilibrium of heterogeneous congestion control
  Accurate study of dynamics
  Conclusion
Outline and Highlight

• Outline
  
  Introduction to congestion control and its current theory
  Equilibrium of heterogeneous congestion control
  Accurate study of dynamics
  Conclusion

• Highlight
  
  Poincare-Hopf index theorem and confirmation with real routers and fibers
  The first accurate stability prediction of TCP/AQM
  General equilibrium theory in economics, Asymmetric matrices in mathematics
Introduction: Congestion Control

- Sources: control data rates according to congestion signals
- Links: adapt congestion signals based on bandwidth utilization
- A purely distributed system!
A Global View of Congestion Control

A distributed feedback control system

- Resource Allocation: Equilibrium (Optimization Theory)
- Interconnected Dynamical System: Dynamics (Control Theory)
A Global View of Congestion Control

A distributed feedback control system

- Resource Allocation: Equilibrium (Optimization Theory)
- Interconnected Dynamical System: Dynamics (Control Theory)

Help answer questions like

- How to understand the whole system for arbitrary (possibly very complex) topology?
- Can all local behaviors of sources and links achieve a globally coordinated goal?
- Are current congestion control schemes scalable? Better ones?

[Kelly-Maulloo-Tan 98], [Low-Lapsley 99], [Mo-Walrand 00]...
Model and Notations

- \( L \) links, indexed by \( l = 1, \ldots, L \). Link \( l \) has a finite capacity \( c_l \) and a congestion signal \( p_l(t) \).

- \( N \) flows, indexed by \( i = 1, \ldots, N \). Flow \( i \) maintains a rate \( x_i(t) \).

- \( f_i \) abstracts the TCP algorithm of flow \( i \).

- \( g_l \) describes the AQM (Active Queue Management) algorithm at link \( l \).

\[
\dot{x}_i = f_i \left( x_i(t), \sum_{l \in L(i)} p_l(t) \right)
\]
\[
\dot{p}_l = g_l \left( \sum_{i:l \in L(i)} x_i(t), p_l(t) \right)
\]
An Example of TCP: TCP Reno

TCP Reno Updating rule (Additive Increase Multiplicative Decrease):

\[
\dot{x}_i = f_i \left( x_i(t), \sum_{l \in L(i)} p_l(t) \right) = \frac{1 - q_i(t)}{\tau_i(t)^2} - \frac{1}{2} q_i(t)x_i^2(t)
\]

\(q_i(t) = \sum_{l \in L(i)} p_l(t)\): packet loss rate;
\(\tau_i(t)\): round trip time (RTT).
TCP Reno Updating rule (Additive Increase Multiplicative Decrease):

\[
\dot{x}_i = f_i \left( x_i(t), \sum_{l \in L(i)} p_l(t) \right) = \frac{1 - q_i(t)}{\tau_i(t)^2} - \frac{1}{2} q_i(t) x_i^2(t)
\]

\[q_i(t) = \sum_{l \in L(i)} p_l(t): \text{packet loss rate;}
\]
\[\tau_i(t): \text{round trip time (RTT).}
\]

At equilibrium:

\[q_i = \frac{2}{2 + \tau_i^2 x_i^2}
\]
TCP Reno Updating rule (Additive Increase Multiplicative Decrease):

\[ \dot{x}_i = f_i \left( x_i(t), \sum_{l \in L(i)} p_l(t) \right) = \frac{1 - q_i(t)}{\tau_i(t)^2} - \frac{1}{2} q_i(t)x_i^2(t) \]

\( q_i(t) = \sum_{l \in L(i)} p_l(t) \): packet loss rate;
\( \tau_i(t) \): round trip time (RTT).

At equilibrium:

\[ q_i = \frac{2}{2 + \tau_i^2 x_i^2} \]

Utility function:

\[ \max_{x_i \geq 0} U_i(x_i) - x_i q_i \]

\[ U_i(x_i) = \frac{\sqrt{2}}{\tau_i} \text{atan} \left( \frac{\tau_i x_i}{\sqrt{2}} \right) \]
Main Theorem: The equilibrium rate vector solves

$$\max_{x \geq 0} \sum_i U_i(x_i) \quad \text{subject to} \quad \sum_{i:l \in L(i)} x_i \leq c_l$$

The source and link protocols serve as a primal-dual algorithm.
Current Theory and Its Impact

Main Theorem: The equilibrium rate vector solves

$$\max_{x \geq 0} \sum_i U_i(x_i) \quad \text{subject to} \quad \sum_{i:l \in L(i)} x_i \leq c_l$$

The source and link protocols serve as a primal-dual algorithm.

Proof: the physical congestion signal ⇔ the mathematical lagrange multipliers.
Main Theorem: The equilibrium rate vector solves

\[
\max_{x \geq 0} \sum_i U_i(x_i) \quad \text{subject to} \quad \sum_{i: l \in L(i)} x_i \leq c_l
\]

The source and link protocols serve as a primal-dual algorithm.

Proof: the physical congestion signal ⇔ the mathematical lagrange multipliers.

Implication, Application and Impact:

Basic properties: existence, uniqueness, optimality

Engineering practice: new protocol design

Wireless networks

Other directions: random arrival, game players . . .
Outline

• Introduction to congestion control and its current theory
• Equilibrium of heterogeneous congestion control
• Accurate study of dynamics
• Conclusion
Outline

- Introduction to congestion control and its current theory

- **Equilibrium of heterogeneous congestion control**
  
  Heterogeneity of congestion signals

  The same congestion signal but different algorithms: **Homogeneous**

- Accurate study of dynamics

- Conclusion
Heterogeneous Congestion Control

- TCP Reno does not work well for networks with large bandwidth-delay products.
  
  Requires extremely small equilibrium loss rate.

  Becomes harder to maintain flow level stability.
Heterogeneous Congestion Control

- TCP Reno does not work well for networks with large bandwidth-delay products.
  
  Requires extremely small equilibrium loss rate.
  
  Becomes harder to maintain flow level stability.

- One solution: use other congestion signals

- Heterogeneous Protocols: Protocols that use different congestion signals
Heterogeneous Congestion Control

- TCP Reno does not work well for networks with large bandwidth-delay products.
  Requires extremely small equilibrium loss rate.
  Becomes harder to maintain flow level stability.

- One solution: use other congestion signals

- Heterogeneous Protocols: Protocols that use different congestion signals

- FAST TCP (queueing delay based):
  At equilibrium,
  \[ x_i = \frac{\alpha_i}{p_i} \]
  Utility function:
  \[ U_i(x_i) = \alpha_i \log x_i \]
Why Study the Heterogeneous Case?

- Various proposals that use different congestion signals
  queueing delay (CARD, DUAL, Vegas, FAST)
  packet loss (Reno and its variants)
  both loss and delay (Westwood, Compound TCP)
  one bit ECN (IETF RFC 2481)
Why Study the Heterogeneous Case?

- Various proposals that use different congestion signals
  queueing delay (CARD, DUAL, Vegas, FAST)
  packet loss (Reno and its variants)
  both loss and delay (Westwood, Compound TCP)
  one bit ECN (IETF RFC 2481)

- The Linux operating system already allows users to choose from a variety of congestion control algorithms since kernel version 2.6.13
Why Study the Heterogeneous Case?

- Various proposals that use different congestion signals
  queueing delay (CARD, DUAL, Vegas, FAST)
  packet loss (Reno and its variants)
  both loss and delay (Westwood, Compound TCP)
  one bit ECN (IETF RFC 2481)

- The Linux operating system already allows users to choose from a variety of congestion control algorithms since kernel version 2.6.13

- Compound TCP which uses multiple congestion signals is part of MS Windows Vista and Windows 2008 TCP stack.
Why Study the Heterogeneous Case?

- Various proposals that use different congestion signals
  - queueing delay (CARD, DUAL, Vegas, FAST)
  - packet loss (Reno and its variants)
  - both loss and delay (Westwood, Compound TCP)
  - one bit ECN (IETF RFC 2481)

- The Linux operating system already allows users to choose from a variety of congestion control algorithms since kernel version 2.6.13

- Compound TCP which uses multiple congestion signals is part of MS Windows Vista and Windows 2008 TCP stack.

- Network will become more heterogeneous.
Why Study the Heterogeneous Case?

- Various proposals that use different congestion signals
  - queueing delay (CARD, DUAL, Vegas, FAST)
  - packet loss (Reno and its variants)
  - both loss and delay (Westwood, Compound TCP)
  - one bit ECN (IETF RFC 2481)

- The Linux operating system already allows users to choose from a variety of congestion control algorithms since kernel version 2.6.13

- Compound TCP which uses multiple congestion signals is part of MS Windows Vista and Windows 2008 TCP stack.

- Network will become more heterogeneous.

- How do we understand it? How can we manage it?
A Motivating Example

Path1 (FAST) → Link1 → Link2 → Link3 → Path2 (FAST)

Path3 (Reno)
A Motivating Example

Equilibrium Rate Allocation:

- Path1 (FAST)
- Path2 (FAST)
- Path3 (Reno)

Diagram showing link connections and rate allocation.
Model

Link $l$: an intrinsic price $p_l$, other “effective prices” $m_i^j(p_l)$.
E.g., $p_l$: queue length, $p^1_l$: loss probability, $p^2_l$: queueing delay.
With RED algorithm:

$$m^1_l(p_l) = \max(1, kp_l) \quad m^2_l(p_l) = \frac{p_l}{c_l}$$
Link \( l \): an intrinsic price \( p_l \), other “effective prices” \( m_i^j(p_l) \).

E.g., \( p_l \): queue length, \( p_l^1 \): loss probability, \( p_l^2 \): queueing delay.

With RED algorithm:

\[
m_i^1(p_l) = \max(1, kp_l) \quad m_i^2(p_l) = \frac{p_l}{c_l}
\]

Homogeneous case:

\[
\dot{x}_i = f_i \left( x_i(t), \sum_{l \in L(i)} p_l(t) \right)
\]

Heterogeneous case:

\[
\dot{x}_i^j = f_i^j \left( x_i^j(t), \sum_{l \in L(j,i)} m_l^j(p_l(t)) \right)
\]
Results

- Existence
- Uniqueness
- Efficiency
- Fairness
- Stability
- Solution: A slow timescale control
Results

- Existence
- **Uniqueness: Number of equilibria**
- Efficiency
- Fairness
- Stability
- Solution: A slow timescale control
Results

- Existence

- Uniqueness: Number of equilibria
  - How many of them? Can the number be 0, 1, 2, 3, $\infty$?
  - Two equilibria and both are stable?
  - A globally unique and stable equilibrium?

- Efficiency

- Fairness

- Stability

- Solution: A slow timescale control
Local Uniqueness is Generic

• multiple locally unique equilibria with different sets of bottleneck links
Local Uniqueness is Generic

- multiple locally unique equilibria with different sets of bottleneck links
- multiple locally unique equilibria with the same set of bottleneck links
Local Uniqueness is Generic

- multiple locally unique equilibria with different sets of bottleneck links
- multiple locally unique equilibria with the same set of bottleneck links
- infinitely many equilibria that are not locally unique
Local Uniqueness is Generic

- multiple locally unique equilibria with different sets of bottleneck links
- multiple locally unique equilibria with the same set of bottleneck links
- infinitely many equilibria that are not locally unique Pathological!

Theorem 1  Given any price mapping functions $m$, any routing matrix $R$ and utility functions $U$,

- for almost all $c$, equilibria are locally unique. Such networks are called regular.
- the number of equilibria for a regular network $(c, m, R, U)$ is finite.

Proof. Use Sard’s theorem.
Equation \( y(p) = c \) (demand equals supply) characterizes equilibrium. 

\[ J(p) = \frac{\partial y(p)}{\partial p}. \]
Equation $y(p) = c$ (demand equals supply) characterizes equilibrium. $J(p) = \partial y(p)/\partial p$.

Define an index $I(p)$ of $p \in E$ (set of equilibrium) as

$$I(p) = \begin{cases} 
1 & \text{if } \det(J(p)) > 0 \\
-1 & \text{if } \det(J(p)) < 0
\end{cases}$$
Global Description: Index Theorem

Equation \( y(p) = c \) (demand equals supply) characterizes equilibrium. \( J(p) = \frac{\partial y(p)}{\partial p} \).

Define an index \( I(p) \) of \( p \in E \) (set of equilibrium) as

\[
I(p) = \begin{cases} 
1 & \text{if } \det(J(p)) > 0 \\
-1 & \text{if } \det(J(p)) < 0
\end{cases}
\]

**Theorem 2** Given any regular network, we have

\[
\sum_{p \in E} I(p) = (-1)^L
\]

where \( L \) is the number of links.

**Proof.** Use Poincare-Hopf index Theorem.
Two Corollaries

Corollary 1  A regular network has an odd number of equilibria.

Corollary 2  If all equilibria are locally stable, then there is exactly one equilibrium.
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty?
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty? Asymmetry!
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty? **Asymmetry**!

Further Development:

\[-J(p) = \sum_{j=1}^{J} R^j \frac{\partial x^j}{\partial q^j}(p) (R^j)^T \frac{\partial m^j}{\partial p}(p)\]
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty? Asymmetry!

Further Development:

\[-\mathbf{J}(p) = \sum_{j=1}^{J} R^j \frac{\partial x^j}{\partial q^j}(p) \left( R^j \right)^T \frac{\partial m^j}{\partial p}(p)\]

- Uniqueness: $\det(\mathbf{J}) > 0$. 
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty? **Asymmetry**!

Further Development:

\[-\mathbf{J}(p) = \sum_{j=1}^{J} R^j \frac{\partial x^j}{\partial q^j}(p) (R^j)^T \frac{\partial m^j}{\partial p}(p)\]

- **Uniqueness**: \(\det(\mathbf{J}) > 0\).
- **Local stability**: All eigenvalues of \(\mathbf{J}\) have positive real parts.
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty? Asymmetry!

Further Development:

\[-J(p) = \sum_{j=1}^{J} R^j \frac{\partial x^j}{\partial q^j}(p) \left( R^j \right)^T \frac{\partial m^j}{\partial p}(p)\]

- Uniqueness: \(\text{det}(J) > 0\).

- Local stability: All eigenvalues of \(J\) have positive real parts.

- Global stability: \(v'Ju > 0\)
From Homogeneity to Heterogeneity

What is the fundamental mathematical difficulty? **Asymmetry!**

Further Development:

\[-J(p) = \sum_{j=1}^{J} R_j \frac{\partial x^j}{\partial q^j}(p) (R^j)^T \frac{\partial m^j}{\partial p}(p)\]

- **Uniqueness:** \(\det(J) > 0\).
- **Local stability:** All eigenvalues of \(J\) have positive real parts.
- **Global stability:** \(v'Ju > 0\)

**\(J = 1, J\) is positive definite.** How about \(J > 1\)? \(J\) becomes asymmetric!
Theorem 3  The equilibrium of a regular network is globally unique and stable if the degree of heterogeneity of price mapping functions is properly bounded.

- protocol independent: uni-protocol case
- link independent: default RED router parameters work
Outline

• Introduction to congestion control and its current theory

• Equilibrium of heterogeneous congestion control

• Accurate study of dynamics

• Conclusion
Outline

• Introduction to congestion control and its current theory

• Equilibrium of heterogeneous congestion control

• Accurate study of dynamics
  make sure the system enjoy desirable equilibrium properties
  avoid under-utilization of link bandwidth
  minimize delay jitters, important for certain real time applications

• Conclusion
Dynamics of Congestion Control

• Equilibrium: Utility Maximization Problem.

• Dynamics:
  
  Local stability: [Johari-Tan 00], [Paganini-Doyle-Low 01], [Hollot-Misra-Towsley-Gong 01], [Vinnicombe 02], [Massoulie 02], [Low-Paganini-Wang-Doyle 03], [Wang-Wei-Low 05] . . .

  Global stability: [Hollot-Chait 01], [Wang-Paganini 02],[Alpcan-Basar 03], [Papachristodoulou-Li-Doyle 04], [Wen-Arcak 04], [Ying-Dullerud-Srikant 06]. . .

• State of the art:
  
  Quantitative results on equilibrium,

  Qualitative study. Cannot compare predictions with packet level simulations quantitatively. Sometimes even worse...
Why Are They Inaccurate

- Window based congestion control

- Consequences:
  - **Burstiness**: faster convergence
  - **Ack-clocking**: scale with RTT, less overshoot

- These microscopic properties matter for dynamics!
Dynamics of Congestion Control

TCP FLOWS

\[ RTT_i(t) = d_i + p(t) \]
Dynamics of Congestion Control

- FAST TCP Window Evolution (Default stepsize $\gamma_i = 0.5$):

$$\dot{w}_i(t) = -\gamma_i \frac{p(t)}{(d_i + p(t))^2} w_i(t) + \gamma_i \frac{\alpha_i}{d_i + p(t)}.$$

- Existing work suggest that FAST TCP is unstable for large enough delay!

- Experiments show that a single FAST TCP flow is always stable regardless of delay!
The integrator link model may lag the true dynamics.
Existing Models

Static link model:

\[
\sum_{i=1}^{N} \frac{w_i(t - \tau_i^f)}{d_i + p(t)} = c.
\]

The static link model may lead the true dynamics.
A New Model

Joint link model:

\[
\int_{t}^{t+d+p(t)} x_i(s) ds = w_i(t + d + p(t)).
\]

The new joint link model is more accurate.
Accurate Closed Loop Validation

The first time in the literature of TCP/AQM!

$c = 10000$ pkt/s. $d_1 = 400$ ms and $d_2 = 700$ ms. Both flows use $\alpha = 50$
Accurate Closed Loop Validation

The first time in the literature of TCP/AQM!

\( c = 10000 \text{ pkt/s. } d_1 = 400 \text{ ms and } d_2 = 700 \text{ ms. Both flows use } \alpha = 50 \)

- Stable
- Unstable
- Integrator link model prediction
- Joint link model prediction
- Static link model prediction

Stepsize

1.23
1.65
1.75
1.83
Accurate Closed Loop Validation

\[\gamma = 1.23\]

“unstable” under integrator link model

\[\gamma = 1.83\]

“stable” under static link model
Accurate Closed Loop Validation

$\gamma = 1.65$

“stable” under joint link model

$\gamma = 1.75$

“unstable” under joint link model
Loop Gain for FAST + New Model

\[
L(s) = \sum_{i=1}^{N} \mu_i L_i(s) = \sum_{i=1}^{N} \mu_i \frac{s + \frac{1}{\tau_i}}{s + \frac{1}{\hat{\tau}}} \frac{d_i \gamma_i e^{-\tau_i s}}{\tau_i^2 s + \gamma_i p}
\]

where

\[
\mu_i = \frac{x_i}{c} = \frac{\alpha_i}{cp}
\]

and

\[
\frac{1}{\hat{\tau}} = \sum_{i=1}^{N} \mu_i \frac{1}{\tau_i}
\]
Theorem 4 If $\gamma \leq M(f(\tau))$, FAST TCP operating over a single link is always locally stable. Here $M(f(\tau))$ only depends on relative heterogeneity of delay distribution, which is independent of absolute values of delays.

- Features:
  
  Directly deal with the sum $L(j\omega)$ instead of individual $L_i(j\omega)$
  
  Bounding region is different for different $\omega$
Summary

• Introduction to the internet congestion control and its current theory

• The current theory breaks down with heterogeneous protocols
  - Interesting behaviors: theoretically and experimentally
  - A new theory: existence, uniqueness, optimality, and stability

• The current theory is not enough to study dynamics quantitatively
  - Work directly with window instead of rate and use a fundamental invariant integral equation to relate both.
  - Provide the first accurate stability prediction
  - Resolve the previous discrepancy between theory and practice