A Bayesian Bandit Approach to Adaptive Delay-based Congestion Control

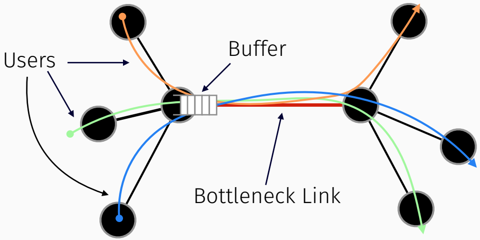
S. D'Aronco and P. Frossard

Packet Video '18, Amsterdam

Signal Processing Laboratory (LTS4)

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Congestion Control Background



- N users share a network link
- x_n : sending rate of user n
- $u_n(x_n)$: utility of user n
- C : capacity of bottleneck link / buffer draining rate

Goal: Maximize overall utility subject to capacity constraints

loss-based CC adapts the rate according to the experienced losses

delay-based CC adapts the rate according to the experienced delay

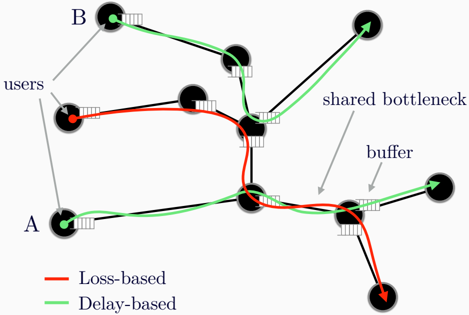
Delay Sensitive Applications

Delay sensitive applications (e.g., VoIP applications) aim at operating at low delay, therefore they must use a delay-based congestion control algorithm.

Delay Sensitive Applications

Delay sensitive applications (e.g., VoIP applications) aim at operating at low delay, therefore they must use a delay-based congestion control algorithm.

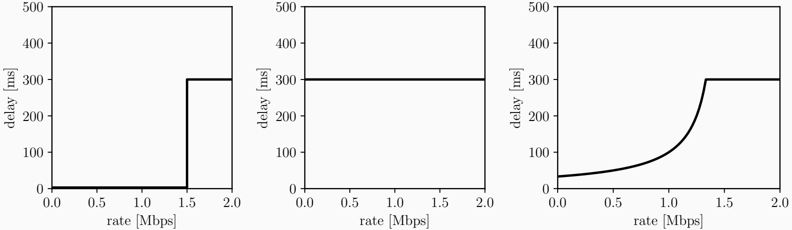
However, delay-based algorithms suffer when competing against loss-based ones (e.g., TCP)



Rate-Delay Tradeoff for Delay-based CC

single link with capacity C and buffer siye Q_{MAX}

Single User　　vs Loss-based　　vs Delay-based

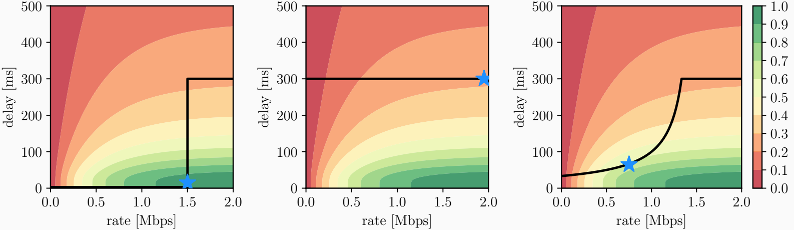


Which equilibrium point should be picked?

Rate-Delay Tradeoff for Delay-based CC

single link with capacity C and buffer siye Q_{MAX}

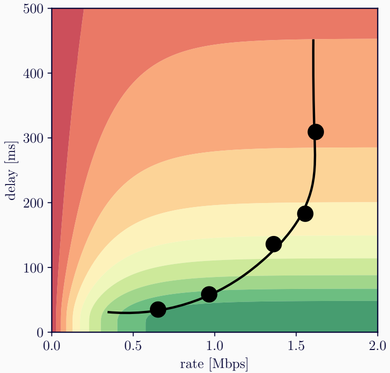
Single User | vs Loss-based | vs Delay-based



Which equilibrium point should be picked?

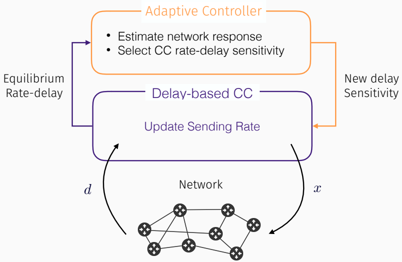
- define utility to model the communication: $u(x, d)$
- pick the equilibrium point that maximizes this metric

Network Sensing



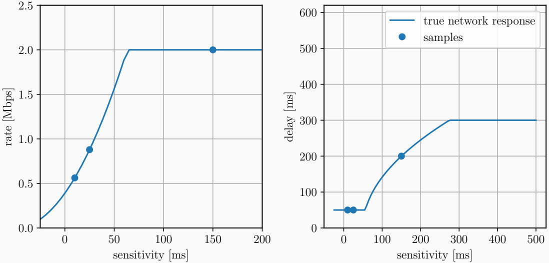
How to sense the network:

- fix sending rate – might degrade network service
- fix delay at equilibrium – might degrade user satisfaction
- fix rate-delay sensitivity by discounting the experience delay by d_{bl} – **good tradeoff**



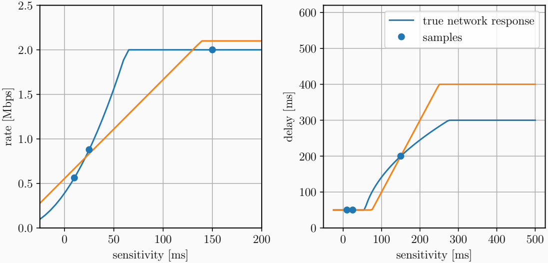
Network Response Model

Goal: infer the network response from point-wise observations



Network Response Model

Goal: infer the network response from point-wise observations

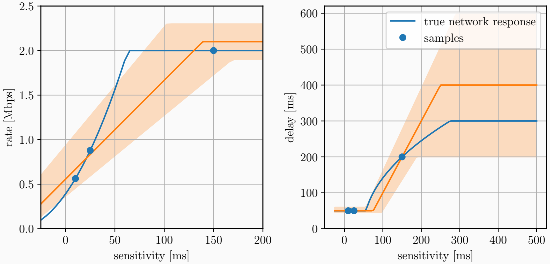


$$x_{\mathsf{eq}} = \min(\theta_x^m d_{\mathsf{bl}} + \theta_x^q, \theta_x^U)$$

$$d_{\mathsf{eq}} = \max(\theta_d^L, \min(\theta_d^m d_{\mathsf{bl}} + \theta_d^q, \theta_d^U))$$

Network Response Model

Goal: infer the network response from point-wise observations



$$x_{\text{eq}} = \min(\theta_x^m d_{\text{bl}} + \theta_x^q, \theta_x^U) \qquad d_{\text{eq}} = \max(\theta_d^L, \min(\theta_d^m d_{\text{bl}} + \theta_d^q, \theta_d^U))$$

By assuming a prior belief on the parameters $\boldsymbol{\theta}_x$ (and $\boldsymbol{\theta}_d$) after observing an equilibrium point, we can infer the posterior belief $p(\boldsymbol{\theta}|x_{\text{obs}}, d_{\text{bl}})$

Bandit Problem

$$\underset{d_{\mathsf{bl}}}{\text{maximize}}\ u(x_{\mathsf{eq}}(d_{\mathsf{bl}};\boldsymbol{\theta}_x), d_{\mathsf{eq}}(d_{\mathsf{bl}};\boldsymbol{\theta}_d))$$

if we know $\boldsymbol{\theta}$

Bandit Problem

$$\underset{d_{\text{bl}}}{\text{maximize}}\ u(x_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_x), d_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_d))$$

if we know θ

$$\underset{d_{\text{bl}}}{\text{maximize}}\ \underbrace{\mathbb{E}_{p(\boldsymbol{\theta})}\left[u(x_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_x), d_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_d))\right]}_{\mathcal{U}(d_{\text{bl}}, p(\boldsymbol{\theta}))}$$

for a single step

Bandit Problem

$$\underset{d_{\text{bl}}}{\text{maximize}}\ u(x_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_x), d_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_d))$$

if we know θ

$$\underset{d_{\text{bl}}}{\text{maximize}}\ \underbrace{\mathbb{E}_{p(\boldsymbol{\theta})}\left[u(x_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_x), d_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_d))\right]}_{\mathcal{U}(d_{\text{bl}},p(\boldsymbol{\theta}))}$$

for a single step

$$\underset{\pi}{\text{maximize}}\ \mathbb{E}_{p\left(\{x^t_{\text{obs}},d^t_{\text{obs}}\}|\{\pi(p^t(\boldsymbol{\theta}))\}\right)}\left[\sum_{t=0}^{\infty}\gamma^t\mathcal{U}\left(\pi(p^t(\boldsymbol{\theta})), p^t(\boldsymbol{\theta})\right)\right]$$

exact problem, too complex

Bandit Problem

$$\underset{d_{\text{bl}}}{\text{maximize}}\ u(x_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_x), d_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_d))$$

if we know θ

$$\underset{d_{\text{bl}}}{\text{maximize}}\ \underbrace{\mathbb{E}_{p(\boldsymbol{\theta})}\left[u(x_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_x), d_{\text{eq}}(d_{\text{bl}};\boldsymbol{\theta}_d))\right]}_{\mathcal{U}(d_{\text{bl}}, p(\boldsymbol{\theta}))}$$

for a single step

$$\underset{\pi}{\text{maximize}}\ \mathbb{E}_{p\left(\{x^t_{\text{obs}}, d^t_{\text{obs}}\}\mid\{\pi(p^t(\boldsymbol{\theta}))\}\right)}\left[\sum_{t=0}^{\infty}\gamma^t\mathcal{U}\left(\pi(p^t(\boldsymbol{\theta})), p^t(\boldsymbol{\theta})\right)\right]$$

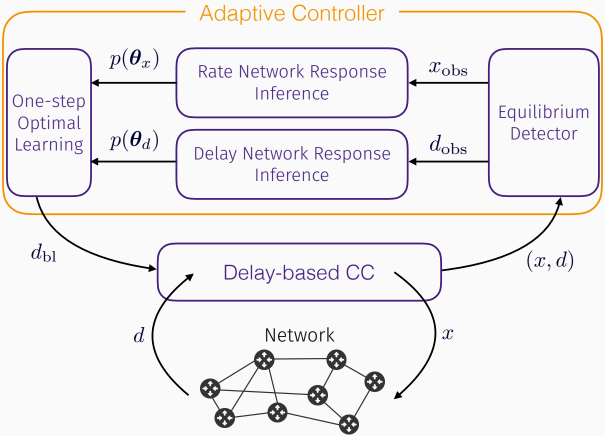
exact problem, too complex

Receding horizon approximation

look one step ahead in the future:

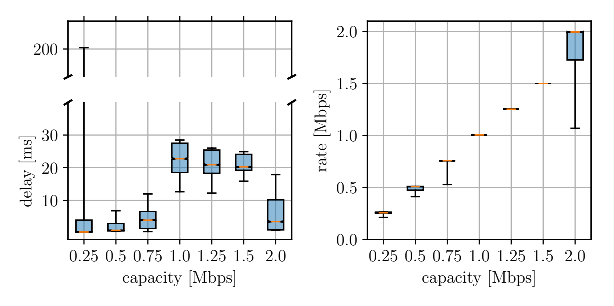
$$\underset{d^t_{\text{bl}}}{\text{maximize}}\ \mathcal{U}(d^t_{\text{bl}}, p^t(\boldsymbol{\theta})) + \frac{\gamma}{1-\gamma}\mathbb{E}_{p^t\left(x_{\text{obs}}, d_{\text{obs}}\mid d^t_{\text{bl}}\right)}\left[\mathcal{U}^\star\left(p^t(\boldsymbol{\theta}\mid x^t_{\text{obs}}, d^t_{\text{obs}}, d^t_{\text{bl}})\right)\right]$$

Implementation



Results - I

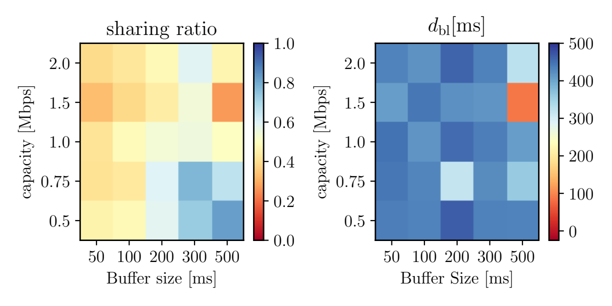
Single user scenario – self inflicted delay for different values of capacity C



the proposed method is able to achieve a low self-inflected delay for wide range of capacity values

Results - II

Competing against a TCP flow - sharing ratio for different values of capacity C and buffer siye Q_{MAX}



the proposed method can reach a fair rate allocation in all the different scenarios

Conclusions

Main advantages:

- model explicitly the communication utility
- estimate a ”global” network response
- take actions to maximize long term utility

Conclusions

Main advantages:

- model explicitly the communication utility
- estimate a ”global” network response
- take actions to maximize long term utility

Future directions/open problems:

- investigate performance in more complex network/scenarios
- investigate theoretical guarantees regarding equilibrium point

Questions?

S. D'Aronco and P. Frossard
stefano.daronco@epfl.ch
Signal Processing Laboratory (LTS4)

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE