On User Strategies in Networks Implementing Congestion Pricing

Tobias Harks
harks@zib.de

Tobias Poschwatta
posch@tkn.tu-berlin.de

Zuse Institute Berlin (ZIB)

TU Berlin

TKN Telecommunication Networks Group

Overview

- Congestion Pricing Model
- Realistic User Strategies
- Download Problem
- Simulation Results
- Wireless Networks

Pricing Model – Elastic Traffic

Sources $s \in S$ have information about:	Links $l \in L$ have information about:
$x_s \sim$ sending rate	$c_l \sim$ link capacity
$U_s(x_s) \sim$ concave utility function	$p_l \sim$ congestion measure (price)
$q_s = \sum_{l \in L(s)} p_l \sim$ path price	$y_l = \sum_{s \in S(l)} x_s \sim$ aggregate transmission rate

If a user bids w_s he receives bandwidth $x_s = \frac{w_s}{q_s}$

Prices p_l are generated by a gradient projection algorithm:

$$\frac{d}{dt} p_l(t) = \gamma [y_l(t) - c_l]^+$$

Optimization Interpretation

User‘s want to maximize their surplus: $\max_{w_s \geq 0} U_s(\frac{w_s}{q_s}) - w_s \quad (1)$

If w_s is continuously varied according to the maximum of (1), the system converges to the unique solution of:

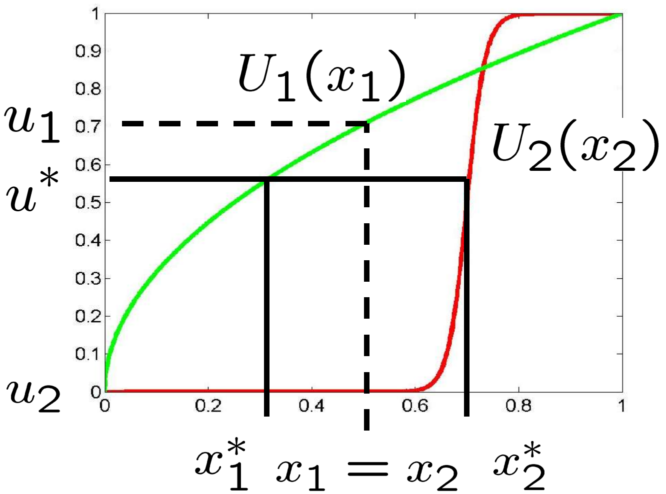
$$(P) \quad \max_{x_s \geq 0} \sum_{s \in S} U_s(x_s)$$
$$s.t. \quad \sum_{s \in S(l)} x_s \leq c_l$$

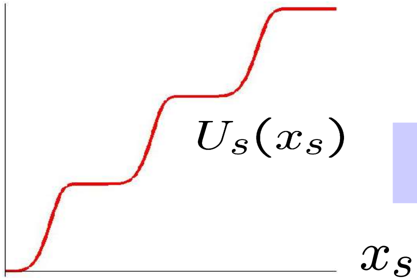
Kelly ’99
TCP-Fast (Low)
Max-Min Kelly Control
(Louginov, 2004 Sigcomm)
Jetmax(L., Infocom 2006)

Real-Time Traffic - Utility Fairness

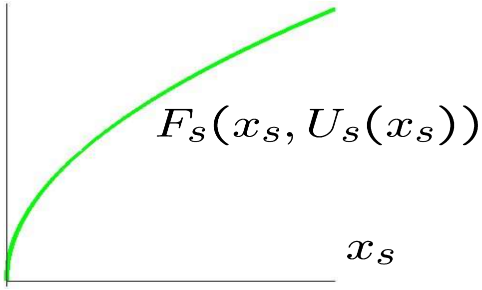
Problem :

- we lose concavity!
- is bandwidth fairness what the user expects?





Transformation



$$(P^*) \quad \max_{x_s \geq 0} \quad \sum_{s\in S} F_s(x_s)$$
$$\text{s.t.} \quad \sum_{s\in S(l)} x_s \leq c_l$$

Realistic User Strategies

- Are user`s preferences captured by static utility functions?

- File download with limited budget and deadline

→ Static utility optimization does not the job!

- Assume user has a budget B_s and wants to download a file of length L_s within time T_s
- we associate with every bid w_s an increasing convex cost function $C_s(w_s)$

Download Problem – Optimal Control

- file length at time t: $l_s(t) = \int_0^t x_s(\tau)d\tau, \quad 0 \le t \le T_s$
- Spend budget at time t: $b_s(t) = \int_0^t w_s(\tau)d\tau, \quad 0 \le t \le T_s$

$$\min_{w_s \in W_s} \int_0^T C_s(w_s(t))dt$$

$$s.t.: \; \dot{l}_s = x_s = \frac{w_s}{q_s}$$

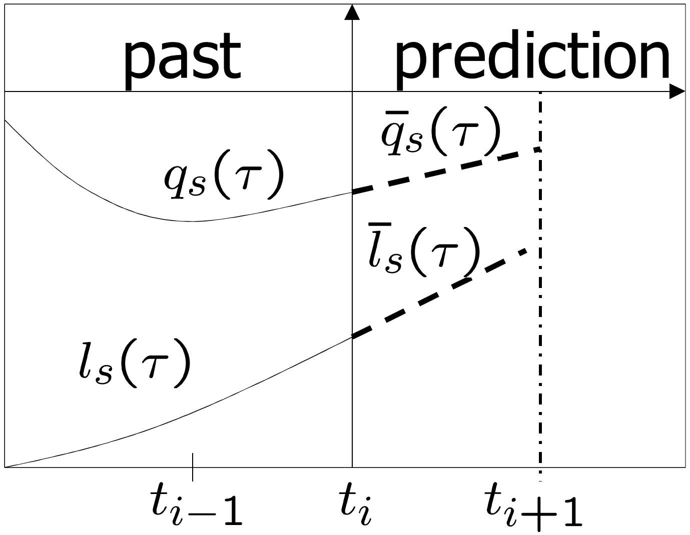
$$\dot{q}_s = f_s(\{x_i | i \text{ uses links of } L(s)\})$$

$$l_s(0) = 0, \; l_s(T_s) = L_s, \; q_s(0) = q_s^0$$

$$b_s(T_s) \le B_s, \; W_s = [w_s^{\min}, w_s^{\max}]$$

Model Predictive Control

- congestion level $q_s(t)$ will fluctuate with level of usage ⇒ system uncertainties!
- online problem!
- Approach: Divide [0,T] in N intervals of length T/N
- Linear Prediction: $\dot{\bar{q}}_s(t) = \zeta_s^i,\ t \in [t_i, t_{i+1}],\ \zeta_s^i \in \mathbb{R}$



Linear Predicting Control

- Solve for every time slot $[t_i, t_{i+1}]$ the following problem:

$$\min_{\bar{w}_s \in W_s} \int_{t_i}^{t_{i+1}} C_s(\bar{w}_s(t))dt$$

$$s.t.: \quad \dot{\bar{l}}_s = \frac{\bar{w}_s}{\bar{q}_s},$$

$$\dot{\bar{q}}_s = \zeta_s^i, \zeta_s^i \in \mathbb{R}$$

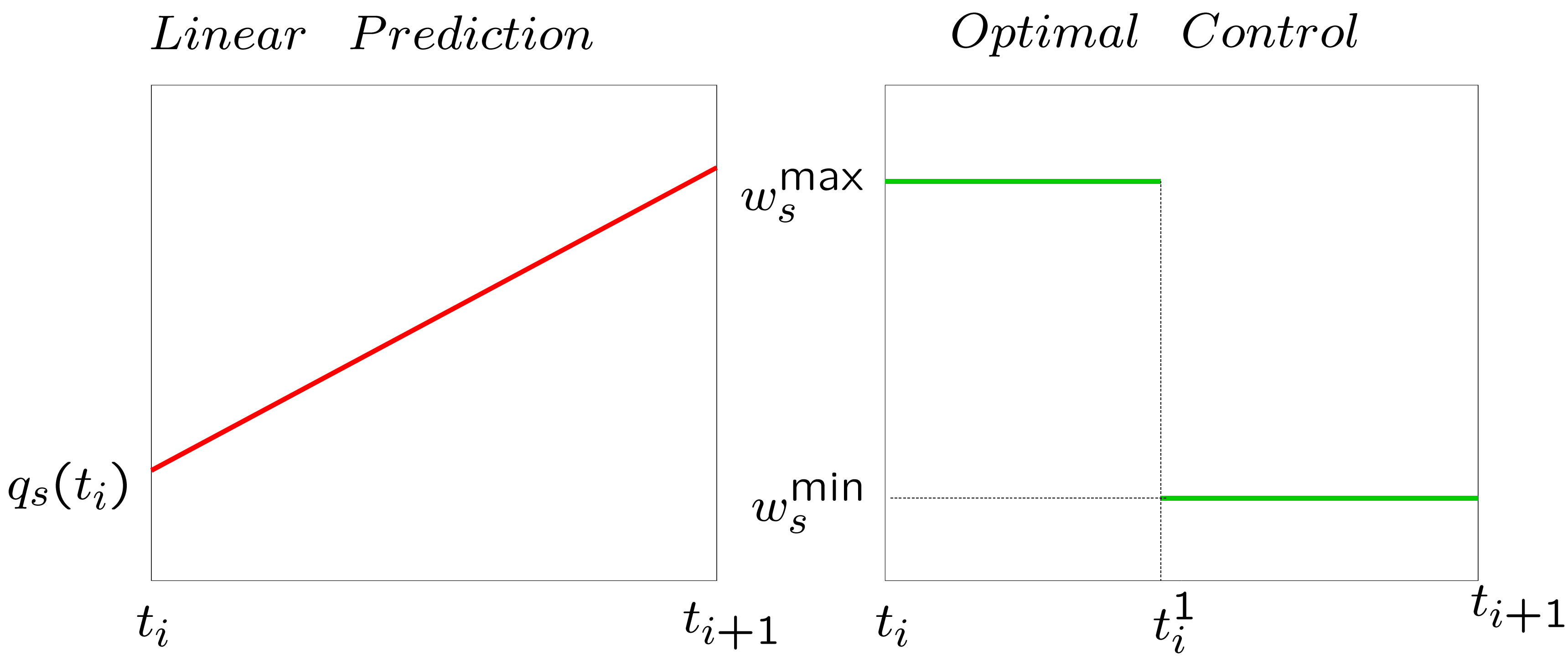
$$\bar{l}_s(t_i) = 0, \quad \bar{l}_s(t_{i+1}) = h\frac{L - l_s(t_i)}{T_s - t_i}$$

$$\bar{q}_s(t_i) = q_s(t_i), \; W_s = [w_s^{\text{min}}, w_s^{\text{max}}]$$

Linear Cost Function

- In the original pricing scheme: $C_s(w_s) = w_s$

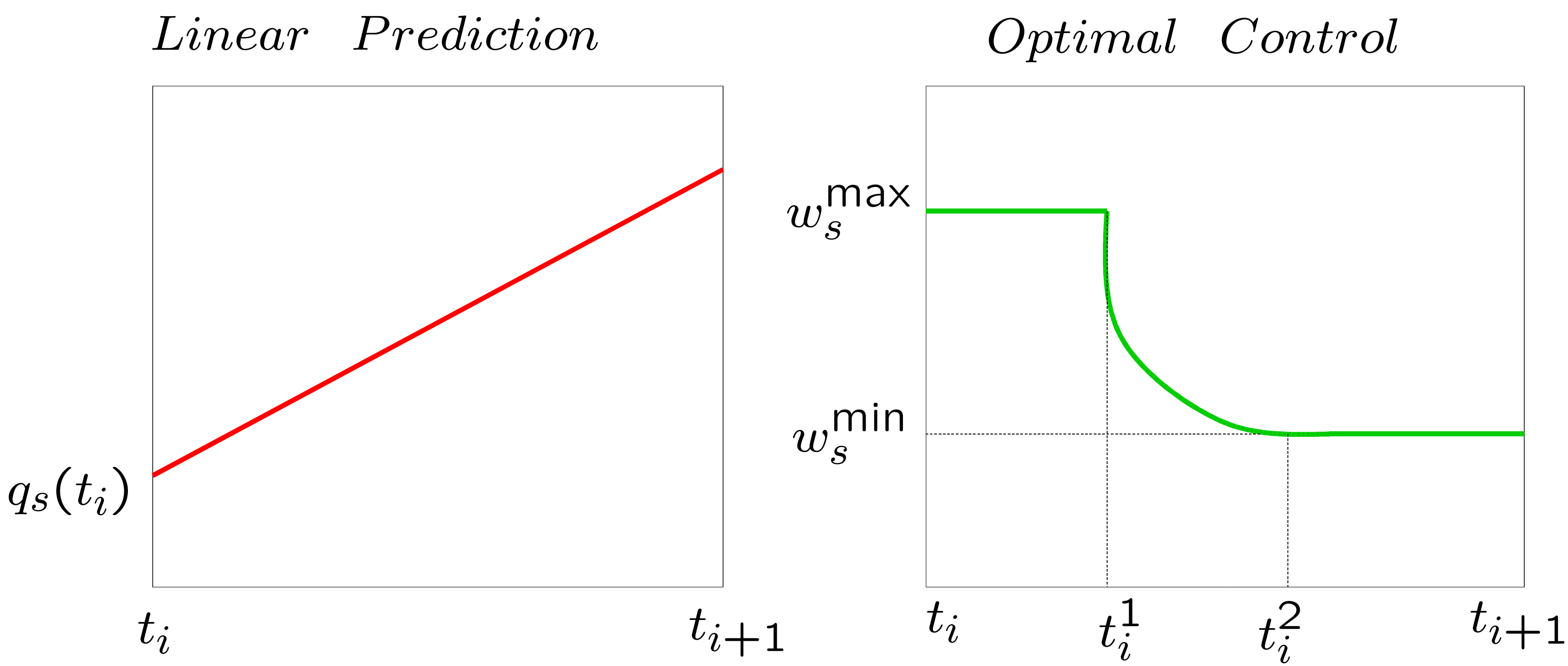
⇒ Optimal control is bang-bang



Convex Cost Function

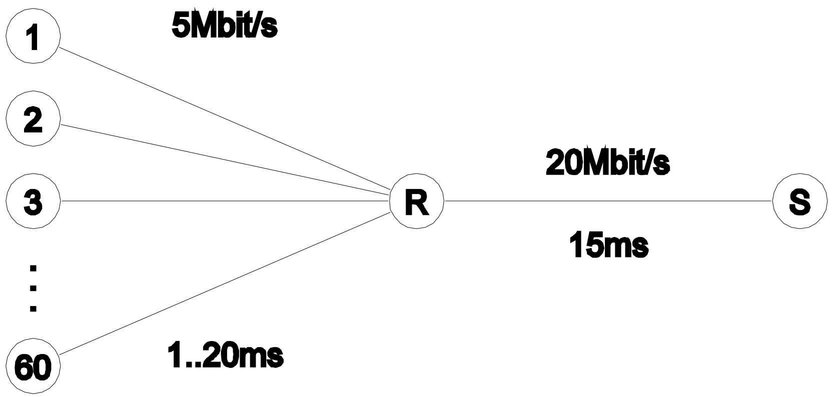
- If the cost function is slightly changed (strictly convex)

$$C_s(w_s) = \frac{w_s^{1+\epsilon}}{1+\epsilon}, \epsilon > 0$$ ⇨ Continuous optimal control

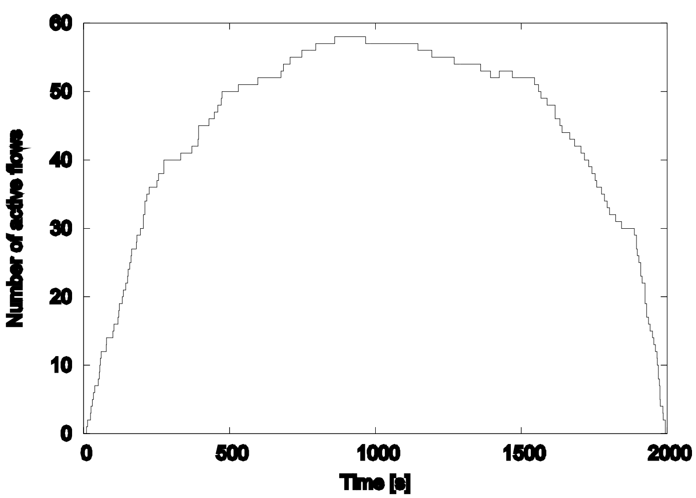


Simulation Setup

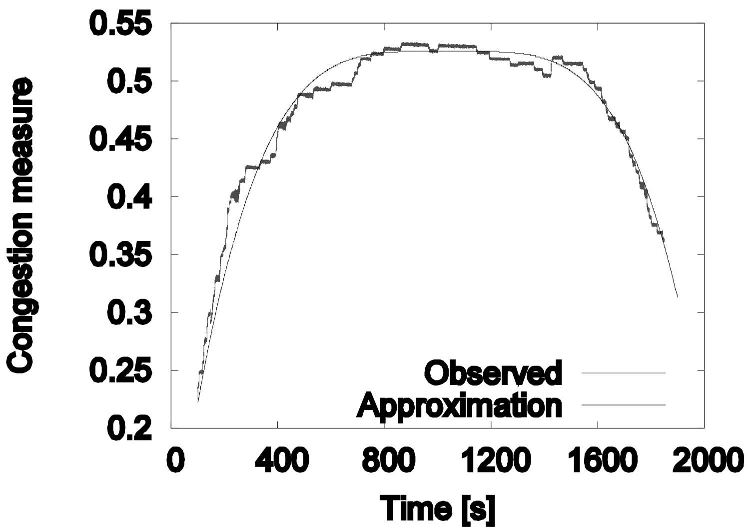
- network topology



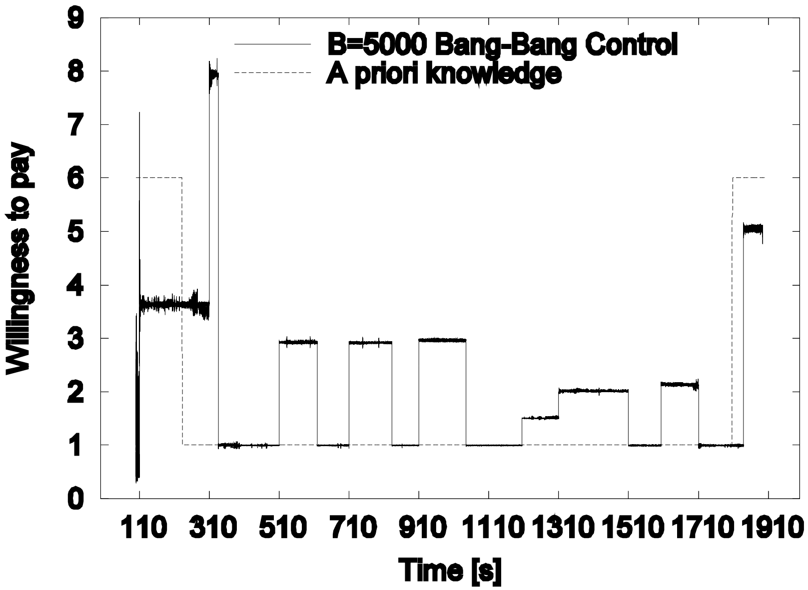
- number of active flows (up to 60)

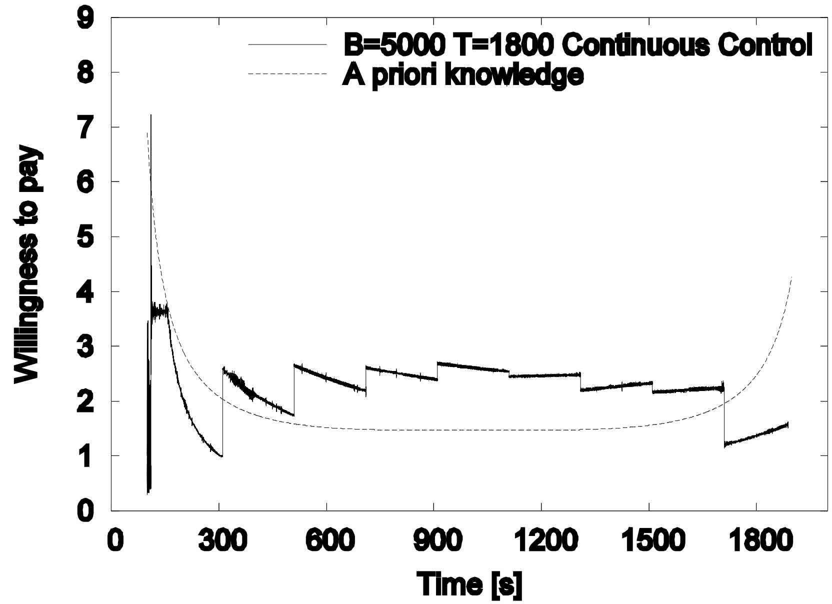


- path price



Bang-Bang and Continuous Control





Significant Improvement for overdimensioned budgets!
Comparison to approach [Gibbens/Kelly ,99]

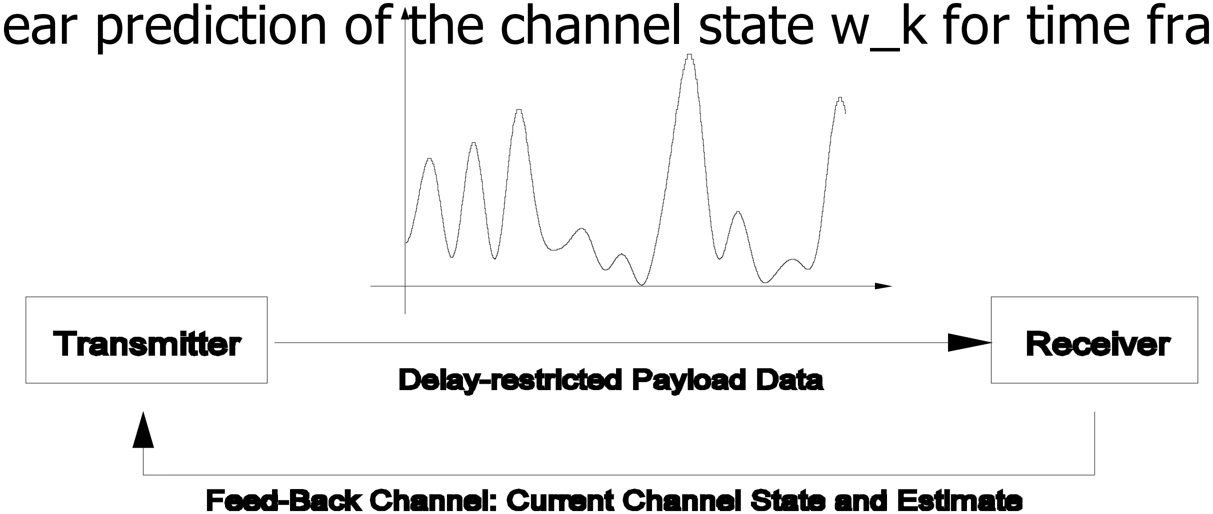
→ 20% Cost savings (B=5000)

Energy Efficient Data Transfer in Wireless Networks (Joint Work with James Gross and Ana Aguiar (TKN))

- Capacity is given by the Shannon capacity formula:

$$C(t) = \log_2(1 + v(t)) = \log_2\left(1 + \frac{p(t)(h(t))^2}{n(t)}\right)$$

- We define the channel state as : $q_k = \frac{h(t_k)^2}{n(t_k)}$
- Problem: Transfer data of length d within time T energy efficiently!
- Dividing the horizon in N time frames
- Linear prediction of the channel state w_k for time frame k



Dynamic Programming Results

Approximation of the cost function for sending data D in frame k by second order polynomial :

$$c_k(w_k, D_k) \approx \frac{aD_k^2 + bD_k + c}{w_k}$$

- This leads to the optimal scheduling problem:

$$\begin{array}{ll} \text{min} & E\left\{\sum_{k\in N} c_k(w_k, D_k)\right\} \\ s.t.: & \sum_{k\in N} D_k = d_1 \end{array}$$

- Optimal policy (data scheduling) is given by:

Optimal Power Control

$$\mu_k(d_k, w_k) = \frac{\prod_{i=k+1}^{N} E_i w_k d_k}{1 + \prod_{i=k+1}^{N} E_i w_k}, \quad \mu_N(d_N, w_N) = d_N$$

$$J_k(d_k, w_k) = \frac{d_k^2 \prod_{i=k+1}^{N} E_i}{1 + w_k \prod_{i=k+1}^{N} E_i}, \quad J_N(d_N, w_N) = \frac{d_N^2}{w_N}$$

- Determine an energyefficient power control strategy in order to transfer D_i within frame i:

$$\min_{p \in [0, P_{Max}]} \int_{t_i}^{t_{i+1}} p(t)dt$$

$$s.t. : \dot{d} = \log_2(1 + pq)$$

$$\dot{q} = \xi_i, \ q(t_i) = q_0^i$$

$$d(t_i) = 0, \ d(t_{i+1}) = D_i,$$

Cross Layer Optimization Projekt

- How can we benefit on the **transport layer** from channel knowledge (attenuation) of the link layer?

  Approach: For the same scenario we solve for every k:

$$\begin{aligned} \min \quad & E\left\{\sum_{k\in N} c_k(w_k, D_k)\right\} \\ s.t.: \quad & \sum_{k\in N} D_k = d_1 \end{aligned}$$

This gives the data D_k^* for every frame (RTT) , and the rate: $E_{rate} = \frac{D_k^*}{\text{Frame Length}}$

Align: E_{rate} with congestion rate (window) x_{TCP}

Cross Layer Optimization Projekt

- How can we benefit on the **transport l… er** from channel knowledge (atte…on) of the …k layer? Appro… the same … for every k:

$\Sigma_{k\in} \ldots _k(\ldots)$

$\Sigma_{k\in} \ldots$

This give… or …, and the rate:

$E \ldots$ …eng…

Align: $I \ldots$ …esti… …d… TCP

The End

Dual Problem - Decentralized Algorithms

Wanted:

Source Algorithm: $x_s(t+1) = H_s(x_s(t), q_s(t))$

Link Algorithm: $p_l(t+1) = G_l(p_l(t), y_l(t))$

Using Lagrangian with Multiplicator:

$$L(x,p) = \sum_{s \in S} U_s(x_s) - p(y-c)$$

Dual Problem:

$$(D) \qquad \min_{p \geq 0} \max_{x \geq 0} L(x,p)$$

Source Algorithm:

$$x_s(t+1) = U_s^{'-1}(q_s(t))$$

Link Algorithm:

$$p_l(t+1) = [p_l(t) + \gamma_l(y_l(t) - c_l)]^+$$

Implemented in REM (Low, 2000) using single bit (ECN)

Results

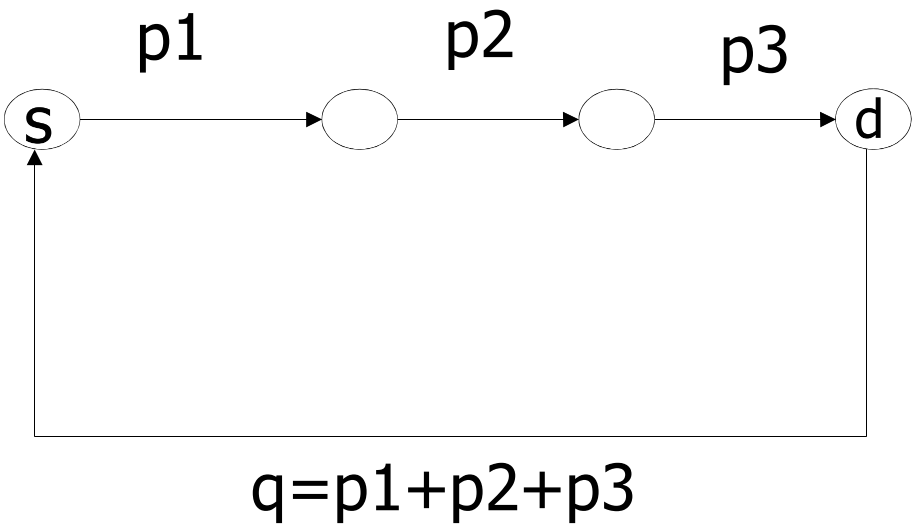
Definition: A rate vector is **utility proportional fair**, if it solves Problem (P*).

Let the rate vector $x = (x_s, s \in S)$ **be utility proportional fair.**

Theorem1: Suppose all users use the same transformation function.
(i) x is **path utility max-min fair**.
(ii) $L(s_1) \subset L(s_2) \Rightarrow U_{s_1}(x_{s_1}) \geq U_{s_2}(x_{s_2})$

Theorem2: Suppose all users use the same transformation function $f_s(q_s, \kappa) = -\sqrt[\kappa]{q_s}, \kappa > 0$ Let the sequence of rate vectors $x(\kappa)$ be utility proportional fair. Then $x(\kappa)$ approaches the **utility max-min fair** rate allocation as $\kappa \to \infty$.

Results



- rate is a function of q
  - q is function of y