# Some Distributed Resource Sharing and Scheduling Problems in Wireless Networks

Yahya Al-Harthi, Dee Denteneer
Seva Shneer, Peter van de Ven, Sem Borst

Alcatel-Lucent Bell Labs & Eindhoven Univ. of Technology

Workshop on Mathematical Modeling and Analysis of Wireless Networks

Toronto, May 8-9, 2008

## Overview

- **Problem I: Linear multi-hop CSMA/CA network**
- **Problem II: Distributed channel-aware scheduling**
- **Problem III: Queue-based channel-aware scheduling in presence of flow-level dynamics**

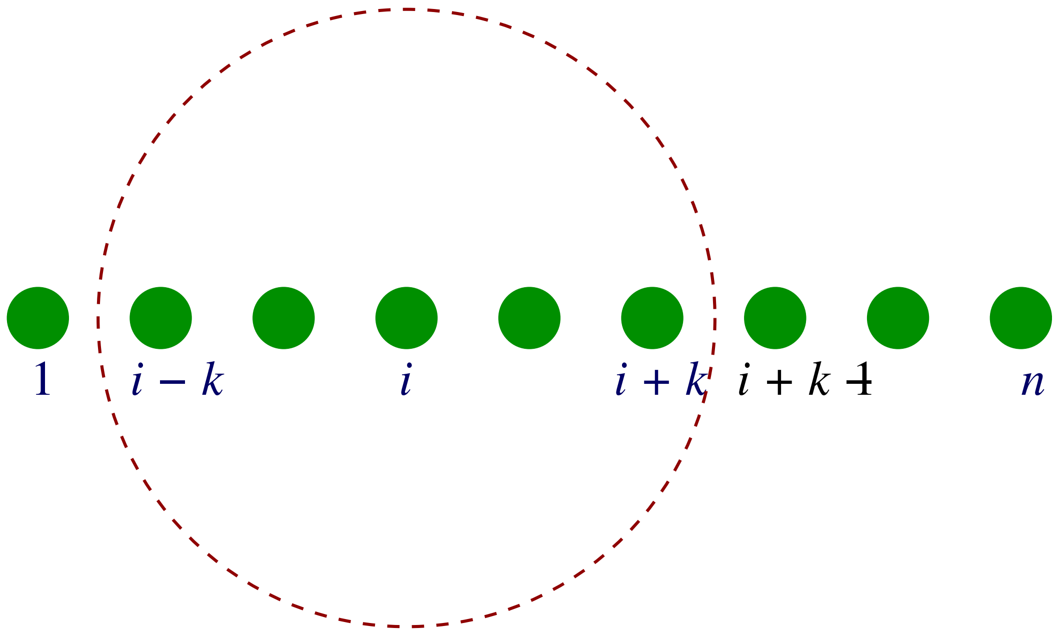
# Problem I: Linear multi-hop CSMA/CA network

Linear multi-hop network with $n$ links, sharing medium according to CSMA/CA type protocol

- $k$-hop interference model, i.e., link can only transmit successfully when no neighboring links within $k$-hop range are active

- Packet transmission times on each link are independent and exponentially distributed with unit mean

- When packets are pending for transmission, and no neighboring links within $k$-hop range are active, link activates at exponential rate $\nu$




**Let $x \in \{0,1\}^n$ represent activity state, with $x_i$ indicating whether or not link $i$ is active**

**Set of feasible activity states**

$$S := \{x \in \{0,1\}^n : \sum_{j=\max\{1,i-k\}}^{\min\{n,i+k\}} x_j \leq 1 \text{ for all } i = 1,\dots,n\}$$

## Scenario A: saturated queues

**First suppose that all links have saturated queues [Wang & Kar (2005), Duvry & Thiran (2006)]**

**Let $X(t) \in S$ represent activity state at time $t$, with $X_i(t)$ indicating whether or not link $i$ is active at time $t$**

**$X(t)$ is reversible Markov process with stationary distribution**

$$\pi(x) \propto \prod_{i=1}^{n} \nu^{x_i} = \nu^{\sum_{i=1}^{n} x_i}, \quad x \in S$$

**Maximum number of simultaneously active links is**

$$m = \max_{x \in S} \sum_{i=1}^{n} x_i = \lceil \frac{n}{k+1} \rceil$$

**Define** $S^* := \{x \in S : \sum_{i=1}^{n} x_i = m\}$ **as set of states with maximum number of active nodes**

**Now consider asymptotic regime where** $\nu \to \infty$

**States** $x \in S^*$ **are asymptotically dominant**

$$\lim_{\nu \to \infty} \pi(x) = \begin{cases} 1/|S^*| & x \in S^* \\ 0 & x \notin S^* \end{cases}$$

**Throughput of link $i$ is $\theta_i = \sum_{x \in S} \pi(x) \mathrm{I}_{\{x_i=1\}}$**

**Limiting throughput of link $i$ is**

$$\theta_i^* = \lim_{\nu \to \infty} \theta_i = \sum_{x \in S^*} \frac{1}{|S^*|} \mathrm{I}_{\{x_i=1\}} = \frac{S_i^*}{|S^*|},$$

**with $S_i^* = \sum_{x \in S^*} \mathrm{I}_{\{x_i=1\}}$**

**Limiting aggregate throughput is**

$$\sum_{i=1}^{n} \theta_i^* = \sum_{i=1}^{n} \lim_{\nu \to \infty} \theta_i = \sum_{x \in S^*} \frac{1}{|S^*|} \sum_{i=1}^{n} \mathrm{I}_{\{x_i=1\}} = m$$

**Observe that $m$ is maximum achievable aggregate throughput, so asymptotically efficient**

**However, tends to be highly unfair**

**Consider case with** $n = 2k + 1$**, then** $|S^*| = k(k+1)/2$**, and** $S_i^* = S_{2k+2-i}^* = k + 1 - i$**,** $i = 1, \ldots, k + 1$

**Thus** $\theta_i^* = \theta_{2k+2-i}^* = \frac{k+1-i}{k(k+1)/2}$

**In particular** $\theta_1^* = \theta_{2k+1}^* = \frac{2}{k+1}$ **while** $\theta_{k+1}^* = 0$**!**

## Scenario B: queue dynamics

**Now consider scenario with queue dynamics and packet transfers**

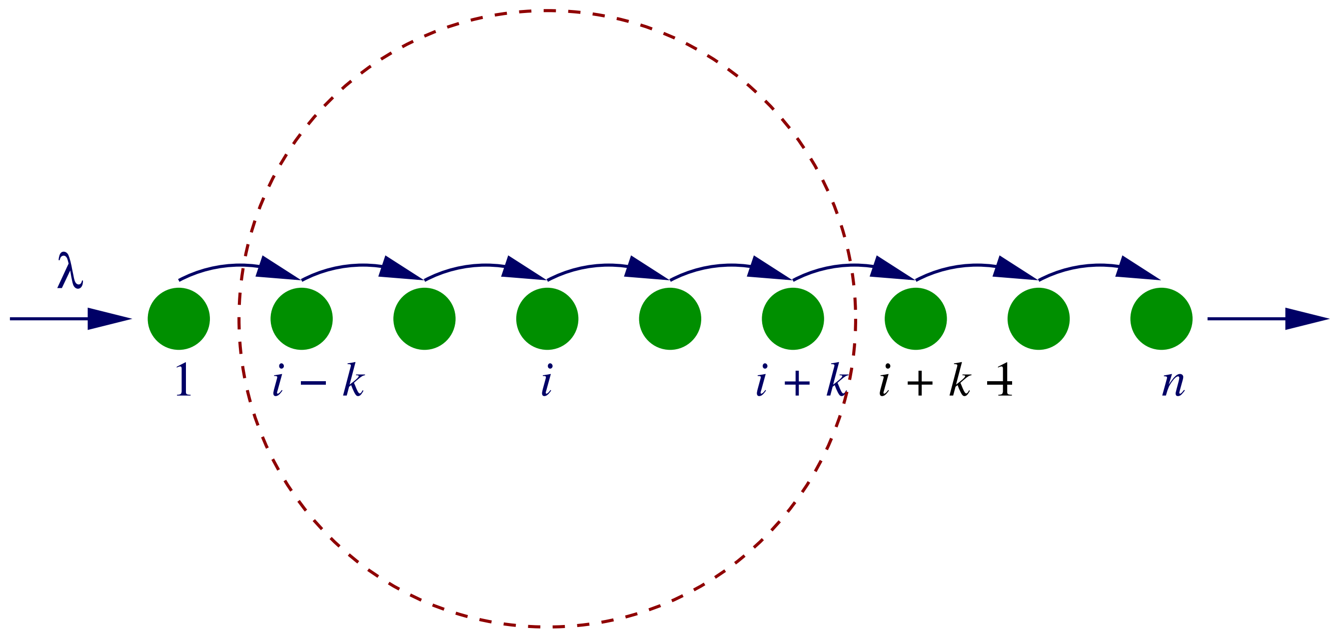
- **Packets offered to link 1 as Poisson process of rate $\lambda$**
- **Packets are forwarded in multi-hop fashion, i.e., link $i$ transfers packets to link $i+1$**
- **Packets leave network after transmission on link $n$**

**Otherwise similar assumptions as before**

- **$k$-hop interference model, i.e., link can only transmit successfully when no neighboring links within $k$-hop range are active**

- **Packet transmission times on each link are independent and exponentially distributed with unit mean**

- **When packets are pending for transmission, and no neighboring links within $k$-hop range are active, link activates at exponential rate $\nu$**




**Let** $x \in \{0,1\}^n$ **represent activity state, and** $y \in \mathbb{N}^n$ **queue size, with** $y_i$ **denoting number of packets in buffer at link** $i$

**Set of feasible states**

$$U = \{(x,y) \in S \times \mathbb{N}^n \text{ with } x_i = 0 \text{ when } y_i = 0\}$$

**Let $X_i(t)$ be 0–1 variable indicating whether or not link $i$ is active at time $t$**

**Let $Y_i(t)$ represent number of packets in buffer of link $i$, either pending for transmission or in process of being transmitted**

$Z(t) = (X_1(t), \ldots, X_n(t), Y_1(t), \ldots, Y_n(t))$ **is Markov process**

**Difficult to analyze...**

**From now on consider asymptotic regime where $\nu \to \infty$**

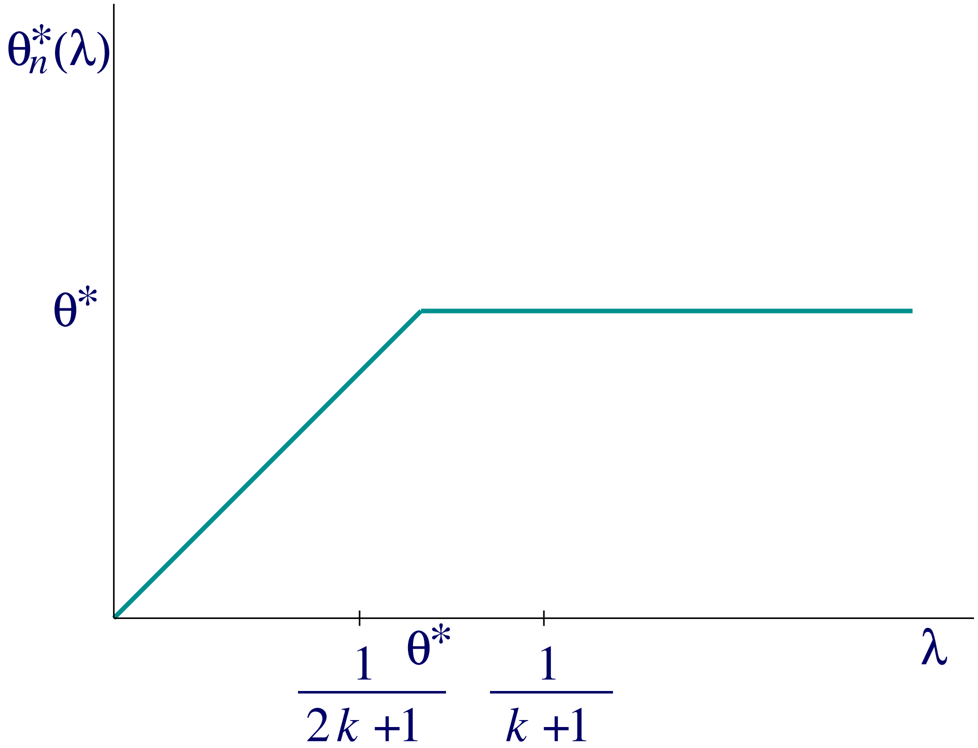
**Denote by $\theta_i^*(\lambda) = \lim_{\nu \to \infty} \theta_i(\lambda)$ limiting throughput of node $i$ as function of $\lambda$**

**Some observations**

- $\theta_1^*(\lambda) \geq \ldots \geq \theta_n^*(\lambda)$
- $\theta_1^*(\lambda) = \ldots = \theta_n^*(\lambda) = \lambda$ **for all values** $\lambda \leq \frac{1}{2k+1}$
- $\theta_n^*(\lambda) \leq \frac{1}{k+1}$ **for all values** $\lambda$

**Define saturation throughput** $\theta^* = \lim_{\lambda \to \infty} \theta_n^*(\lambda)$

**One might expect that** $\theta_n^*(\lambda) = \min\{\lambda, \theta^*\}$ **with** $\theta^* \in \left(\frac{1}{2k+1}, \frac{1}{k+1}\right)$



Let us focus on network with $k = 1$ and $n = 3$, and determine saturation throughput $\theta^* = \lim_{\lambda \to \infty} \theta_3^*(\lambda)$

Link 1 will be saturated for sufficiently large values of $\lambda$

It is easily seen that link 2 will then be saturated as well, since its activity periods are (stochastically) shorter than those of link 1 due to additional competition from link 3

Further observe that $X_2(t) = 1 - X_1(t)$ and $X_3(t) = X_1(t) \mathrm{I}_{\{Y_3(t) \geq 1\}}$

Markov process $Z(t) = (X_1(t), X_2(t), X_3(t), Y_1(t), Y_2(t), Y_3(t))$ may be reduced to $(X_1(t), Y_3(t))$, with

- $X_1(t) \in \{0, 1\}$ indicating whether or not link 1 is active

- $Y_3(t)$ denoting number of packets in buffer of link 3

**Stationary distribution** $\pi(x, y) = \lim_{t \to \infty} \mathrm{P}\{(X_1(t), Y_3(t)) = (x, y)\}$ **is**

$$\pi(0, y) = \frac{1}{5}\left(\frac{1}{3}\right)^{y}, \qquad y \geq 0,$$

$$\pi(1, 0) = \frac{2}{5},$$

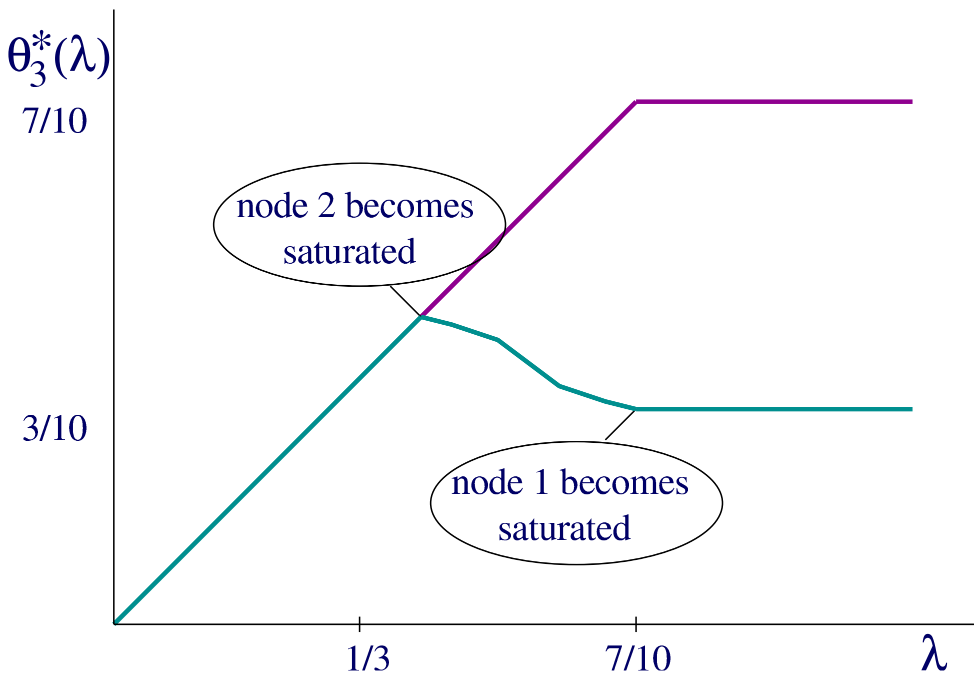
**and**

$$\pi(1, y) = \frac{1}{5}\left(\frac{1}{3}\right)^{y-1}, \qquad y \geq 1$$

**End-to-end system throughput may be obtained as fraction of time that link 3 is active:** $\theta^* = \sum_{y=1}^{\infty} \pi(1, y) = 3/10$

**Equivalently, system throughput may be obtained as fraction of time that link 2 is active:** $\theta^* = \sum_{y=0}^{\infty} \pi(0, y) = 3/10$

**Recall that** $\theta_3^*(\lambda) = \lambda$ **for all** $\lambda \leq 1/3$**!**



**Now consider network with one extra link: $n = 4$, $k = 1$**

**One would expect end-to-end throughput to drop or remain constant compared to $n = 3$, $k = 1$**

**It turns out however that end-to-end throughput in fact goes up**

**Interpretation: bottleneck link 2 benefits from presence of link 4 in competition with links 1 and 3**

**Conjecture:** $\frac{1}{2k+2} \leq \theta^* \leq \frac{1}{2k+1}$ **for all** $n \geq 2k+1$

**In contrast, end-to-end throughput** $\frac{1}{k+1}$ **can be achieved using TDMA type scheme**

# Problem II: Distributed channel-aware scheduling

Multi-access channel with $N$ users, operating in time-slotted fashion

$R_i(t)$ is feasible transmission rate of user $i$ in time slot $t$

$\{R_i(t)\}_{t \geq 1}$ is stationary ergodic sequence

$F_i(r) = \mathrm{P}\{R_i \leq r\}$ continuous distribution function with density $f_i(r) = \mathrm{d}F(r)/\mathrm{d}r$

When several users decide to transmit in same time slot, collision occurs and all transmissions fail

When $R_i(t) = r$, user $i$ transmits in time slot $t$ with probability $P_i(r)$, regardless of decisions of other users

**Pareto-optimal throughput vectors achieved by threshold strategies [Adireddy & Tong (2002), Qin & Berry (2003)]:**

$$P_i(r) = \begin{cases} 0 & r < \gamma_i \\ 1 & r \geq \gamma_i \end{cases}$$

**for some threshold parameter $\gamma_i$, i.e., user $i$ transmits in time slot $t$ if and only if $R_i(t) \geq \gamma_i$**

**For given throughput utilitity function, is it possible to determine optimal transmission thresholds $\gamma_1^*, \ldots, \gamma_N^*$ in distributed fashion without explicit knowledge of rate statistics?**

**Denote expected throughput of user $i$ as function of vector of transmission thresholds $\gamma = (\gamma_1, \ldots, \gamma_N)$**

$$T_i(\gamma) = S_i(\gamma_i) \prod_{j \neq i} F_j(\gamma_j)$$

**with $S_i(\gamma_i) = \mathrm{E}\{R_i \mathrm{I}_{\{R_i > \gamma_i\}}\}$**

## Weighted Proportional Fairness

**Consider aggregate throughput utility function**

$$G(\gamma) = \sum_{i=1}^{N} w_i \log(T_i(\gamma))$$

**Note that**

$$\begin{aligned} G(\gamma) &= \sum_{i=1}^{N} w_i \log\left(S_i(\gamma_i) \prod_{j \neq i} F_j(\gamma_j)\right) \\ &= \sum_{i=1}^{N} w_i \left[\log\left(S_i(\gamma_i)\right) + \sum_{j \neq i} \log(F_j(\gamma_j)\right] \\ &= \sum_{i=1}^{N} \left[w_i \log\left(S_i(\gamma_i)\right) + w_{-i} \log(F_i(\gamma_i)\right], \end{aligned}$$

**with** $w_{-i} := \sum_{j \neq i} w_j$

We obtain

$$\begin{aligned}\frac{\partial G(\gamma)}{\partial \gamma_i} &= \frac{\mathsf{d}}{\mathsf{d}\gamma_i}\left[w_i \log\left(S_i(\gamma_i)\right) + w_{-i}\log(F_i(\gamma_i))\right] \\ &= \frac{w_i}{S_i(\gamma_i)}\frac{\mathsf{d}S_i(\gamma_i)}{\mathsf{d}\gamma_i} - \frac{w_{-i}}{F_i(\gamma_i)}\frac{\mathsf{d}F_i(\gamma_i)}{\mathsf{d}\gamma_i} \\ &= \frac{w_i}{S_i(\gamma_i)}\gamma_i f_i(\gamma_i) - \frac{w_{-i}}{F_i(\gamma_i)} f_i(\gamma_i) \\ &= \left[\frac{w_i\gamma_i}{S_i(\gamma_i)} - \frac{w_{-i}}{F_i(\gamma_i)}\right] f_i(\gamma_i)\end{aligned}$$

- $\frac{w_i\gamma_i f_i(\gamma_i)}{S_i(\gamma_i)}$: marginal relative throughput gain for user $i$

- $\frac{w_{-i} f_i(\gamma_i)}{F_i(\gamma_i)}$: marginal aggregate relative throughput loss for all other users

Thus $\frac{\partial G(\gamma)}{\partial \gamma_i} = 0$ requires

$$w_i \gamma_i F_i(\gamma_i) = w_{-i} S_i(\gamma_i)$$

Above equation may be solved by each individual user estimating $F_i(\gamma_i)$ and $S_i(\gamma_i)$, and either increasing or decreasing transmission threshold $\gamma_i$ by small amount, depending on whether

$$w_i \gamma_i \widehat{F}_i(\gamma_i) < w_{-i} \widehat{S}_i(\gamma_i)$$

or

$$w_i \gamma_i \widehat{F}_i(\gamma_i) > w_{-i} \widehat{S}_i(\gamma_i)$$

## Queue stability

**Now suppose user $i$ generates average of $\lambda_i$ bits per slot**

**There exists transmission strategy that achieves stability if and only if $(\lambda_1, \ldots, \lambda_N) \in \mathcal{T}$, with**

$$\mathcal{T} = \{(T_1(\gamma), \ldots, T_N(\gamma)) : \gamma \in \mathbb{R}_+^N\}$$

**representing achievable throughput region**

**Is it possible to determine suitable transmission thresholds $\gamma_1, \ldots, \gamma_N$ in distributed fashion without explicit knowledge of $(\lambda_1, \ldots, \lambda_N)$ and rate statistics?**

**Approach: periodically update transmission thresholds, using queue lengths as weight factors**

**Either increase or decrease transmission threshold $\gamma_i(t)$ at time $t$ by small amount, depending on whether**

$$w_i(t)\gamma_i(t)\widehat{F}_i(\gamma_i(t)) < w_{-i}(t)\widehat{S}_i(\gamma_i(t))$$

**or**

$$w_i(t)\gamma_i(t)\widehat{F}_i(\gamma_i(t)) > w_{-i}(t)\widehat{S}_i(\gamma_i(t))$$

**with** $w_i(t) = Q_i(t)$ **and** $w_{-i}(t) = \sum_{j \neq i} Q_j(t)$

## Extensions

- **Capture effects, multi-user reception, network setting**
- **General concave utility functions**

## Problem III: Queue-based channel-aware scheduling in presence of flow-level dynamics

**Broadcast channel, operating in time-slotted fashion**

**In each time slot, exactly one flow is selected for transmission**

$R_i(t)$ **is feasible rate of** $i$**-th flow in time slot** $t$**, if selected for transmission**

$Q_i(t)$ **is backlog of** $i$**-th flow in time slot** $t$

## Scenario A: static flow population, packet arrivals

**Static population of $N$ flows**

**Each flow generates bits over time**

**Queue-based scheduling: time slot $t$ is assigned to flow** $i^*(t) := \arg\max_{i=1,\dots,N} Q_i(t) R_i(t)$

**Achieves maximum stability**
**[Tassiulas & Ephremides (1992), McKeown et al. (1996), Neely et al. (2003), Stolyar (2004)]**

# Scenario B: dynamic flow configuration

**Flows arrive over time, at average rate of $\alpha$ per slot**

**Each flow either produces instantaneous traffic burst upon arrival or generates bits over some finite random time period**

**Denote by $B$ total service requirement (in bits) of arbitrary flow**

**Denote by $R_{\text{max}}$ maximum feasible transmission rate**

**Suppose any flow could be served at rate $R_{\text{max}}$ in any time slot**

**Expected number of required transmissions per slot**

$$\rho = \alpha \mathbf{E}\{\lceil B/R_{\text{max}} \rceil\}$$

**Necessary stability condition: $\rho < 1$**

**Utility-based scheduling: time slot $t$ is assigned to flow** $i^*(t) := \arg \max_{i:Q_i(t)>0} w_i(t) R_i(t)$

**Weights $w_i(t)$ set according to $\alpha$-fair utility optimization**

**Achieves maximum stability ($\rho < 1$)**

Note: utility-based strategies in general fail to ensure maximum packet-level stability in case of static flow population

Utility-based strategies do not respond to queue build-ups, suggesting that queue-based strategies could provide potentially better packet-level performance

It turns out however that queue-based strategies may in fact cause instability

Interpretation: flows with large backlogs may receive priority over flows with smaller backlogs, even when their transmission rates are relatively low

In case of static flow population, flows with smaller backlogs will build up large backlogs as well, and situation cannot persist

In presence of flow-level dynamics however, instability manifests itself in form of large number of flows rather than large backlogs per flow

For example, assume

- service requirement $B$ is an (even) constant
- feasible transmission rate $R$ is either 1 or 2 with equal probability

so that $\rho = \alpha B/2$

**Number of flows with residual size larger than $\frac{3}{8}B$ cannot grow without bound as long as $\rho < 1$**

**Suppose that were not true, then flows of residual size in $[\frac{3}{8}B, \frac{3}{4}B]$ can only be served at rate 2**

**Hence flows with residual size smaller than $\frac{3}{8}B$ will not be served at all**

**Thus total fraction of time spent on service of flows with residual size larger than $\frac{3}{8}B$ is at most**

$$\alpha[\frac{3}{8}B/2 + \frac{1}{4}B] = \frac{7}{16}\alpha B = \frac{7}{8}\rho < \frac{7}{8}$$

**Let $p$ be probability that there is at least one flow of residual size larger than $\frac{3}{8}B$, but no such flow with feasible rate 2**

**Probability $p$ is bounded away from zero as long as $\rho < 1$**

**Hence average service rate is bounded away from 2 as long as $\rho < 1$**

**It follows that system is unstable for $\rho > 1 - \epsilon$ for some $\epsilon > 0$**