Scaling laws and medium access

Devavrat Shah
Laboratory for Information and Decision Systems
Department of EECS
Massachusetts Institute of Technology

Based on joint works with

Piyush Gupta (Bell Labs)  Urs Niesen (Bell Labs)  Jinwoo Shin (MIT)
Background

- Complex networks that are engineered
  - Examples:
    - Snail-mail, WWW, P2P, Facebook, Twitter, ...
  - Backbone: a communication network
    - China Post, Telephone, Electronic, Optical, Wireless, ...
Background

- Complex networks that are engineered
  - Examples:
    - Snail-mail, WWW, P2P, Facebook, Twitter, ...
  - Backbone: a communication network
    - *China Post*, Telephone, Electronic, Optical, Wireless, ...

- An important challenge going forward
  - Architecting high-aggregate bandwidth communication network
    - Data centers, access networks, ...
Background

- Access networks
  - Wireless technology provides architecture of choice
  - Examples include
    - Meraki Networks, Rice Network in Houston, ...
Background

- Access networks
  - Wireless technology provides architecture of choice
  - Examples include
    - Meraki Networks, Rice Network in Houston, ...

- Intellectually, there are two road blocks
  - What is the network capacity?
  - How to build capacity achieving networks practically?
Background

- Access networks
  - Wireless technology provides architecture of choice
  - Examples include
    - Meraki Networks, Rice Network in Houston, ...

- Intellectually, there are two road blocks
  - What is the network capacity ?
    - Scaling laws
  - How to build capacity achieving networks practically ?
    - Medium access
Rest of the talk

- Scaling laws for *homogeneous* networks
  - Capacity by means of appropriate medium access

- Medium access
  - Efficient protocol based on distributed random access

- Going forward
  - Scaling laws for heterogenous networks
  - Distributed medium access beyond interference management
Wireless network

- Point-to-point communication
  - Reasonably understood over the past 60 years
    - since Shannon (1948)
Wireless network

- Multi-terminal communication
  - Interference management
Wireless network

- Multi-terminal communication
  - Interference management
  - Co-operation gain
Wireless network

- Multi-terminal communication
  - Interference management
  - Co-operation gain

- Capacity region
  - Requires delicate balance of interference mgnt. & co-operation
    - which is quite challenging
    - not known even for (arbitrary) 3 node network
Wireless network

• Multi-terminal communication
  ○ Interference management
  ○ Co-operation gain

• Capacity region: scaling laws
  ○ Approximation (or pragmatic approach)
  ○ Even for wired networks *nicest* answers
    ▪ are either in form of non-intuitive linear program
    ▪ or, intuitive approximate (spectral) form
Scaling laws: brief history

- Scaling laws for wireless networks
  - Introduced by Gupta and Kumar (2000)
  - Question of interest
    - how does capacity region ‘scale’ with network size?
Scaling laws: brief history

- Result of Gupta and Kumar (2000)
  - Node placement: regular
    - $n$ nodes placed uniformly at random in square of area $n$
Scaling laws: brief history

- Result of Gupta and Kumar (2000)
  - Node placement: regular
  - Communication: protocol model
    - interference treated as noise and no co-operation
    - nearby nodes cannot transmit simultaneously
Scaling laws: brief history

- Result of Gupta and Kumar (2000)
  - Node placement: regular
  - Communication: protocol model
  - Traffic demand: uniform
    - randomly selected $n$ distinct source-destination pairs
Scaling laws: brief history

• Result of Gupta and Kumar (2000)
  ○ Node placement: regular
  ○ Communication: protocol model
  ○ Traffic demand: uniform
  ○ Then, maximal rate achievable per node
    • scales essentially as $\Theta\left(\frac{1}{\sqrt{n}}\right)$
Scaling laws: brief history

- Information theoretic limits on scaling law
  - Node placement: regular
  - Communication: Gaussian fading channel
    - $Y = HX + Z$ with
    - independent fading coefficients $H$. with
    - signal strength attenuating with distance polynomially
      - as per exponent $\alpha > 2$
    - and noise $Z$. being i.i.d. normal (complex) Gaussian
  - Traffic demand: uniform
Scaling laws: brief history

- Information theoretic limits on scaling law
  - Node placement: regular
  - Communication: Physical model in 2-D
    - derived based on Electro-Magnetic wave propagation
    - and additive Gaussian noise
  - Traffic demand: uniform
Scaling laws: brief history

- Information theoretic limits on scaling law
  - Node placement: regular
  - Traffic demand: uniform
  - Communication:
    - Gaussian fading with $\alpha > 3$
    - Or, physical model in 2-D
  - Then, maximal rate achievable per node
    - can scale at most $O\left(\frac{1}{\sqrt{n}}\right)$
Scaling laws: brief history

- Information theoretic limits on scaling law
  - Node placement: regular
  - Traffic demand: uniform
  - Communication:
    - Gaussian fading with $\alpha > 3$
    - Or, physical model in 2-D
  - Then, maximal rate achievable per node
    - can scale at most $O\left(\frac{1}{\sqrt{n}}\right)$

- These results are due to: Gaussian fading channel
Scaling laws: brief history

- Information theoretic limits on scaling law
  - Node placement: regular
  - Traffic demand: uniform
  - Communication:
    - Gaussian fading with $\alpha > 3$
    - Or, physical model in 2-D
  - Then, maximal rate achievable per node
    - can scale at most $O\left(1/\sqrt{n}\right)$

- These results are due to: Physical model in 2-D
Scaling laws: brief history

• Information theoretic limits on scaling law
  ○ Node placement: regular
  ○ Traffic demand: uniform
  ○ Communication:
    - Gaussian fading with $\alpha \in (2, 3)$
  ○ Then, maximal rate achievable per node
    - scales essentially as $\Theta\left( n^{1-\frac{\alpha}{2}} \right)$
    - requires sophisticated network-wide co-operation

• Due to
  ○ Ozgur-Leveque-Tse (2008) and Aeron-Saligrama (2007)
Scaling laws: our result

- Setup in Niesen, Gupta and Shah (2009)
  - Node placement: regular
  - Traffic demand: arbitrary
    - demand vector \( \lambda = [\lambda_{ij}] \in \mathbb{R}^{n \times n} \)
    - earlier setup equiv. to finding \( n\rho \) where \( \lambda_{ij} = \rho, \ \forall \ i, j \)
    - interest in \( \Lambda \), the set of all feasible \( \lambda \)
  - Communication: information theoretic
    - Gaussian fading with \( \alpha > 2 \)
    - Or, physical model in 2-D
Wireless network = A wired tree network

An equivalence in terms of (scaling) capacity region

Under the setup described

Scaling laws: our result
Scaling laws: our result

- Under the setup described
  - An equivalence in terms of (scaling) capacity region
    - Wireless network = A wired tree network
Scaling laws: our result

- Under the setup described
  - An equivalence in terms of (scaling) capacity region
    - Wireless network = A wired tree network
  - The tree is realized using physical layer such that
    - no co-operation, only interference management
      - Gaussian fading with $\alpha > 3$ or Physical 2-D
    - sophisticated network-wide co-operation
      - Gaussian fading with $\alpha \in (2, 3)$
Scaling laws: our result

- Implications of equivalence:
  - Operationally
    - Network layer utilizes tree for routing
    - Wireless or Physical layer is oblivious to demand
Scaling laws: our result

- Implications of equivalence:
  - Characterization of capacity region
    - By means of 2n inequalities
    - That can be easily evaluated
Scaling laws: our result

- Further equivalence:
  - Multicast capacity region (upto scaling)
    - Wireless network = wireline tree network
Summary, thus far

• We started with two questions
  ○ What is the network capacity?
  ○ How to build capacity achieving networks practically?

• Scaling laws answer first question for regular networks

• To practically realize such capacity scaling
  ○ Need simple, distributed algorithms for
    ▪ network-wide interference management
      ▪ for Gaussian fading with $\alpha > 3$ or Physical model
      ▪ and, sophisticated network-wide co-operation
        ▪ for Guassian fading with $\alpha \in (2, 3)$

• Next, distributed medium access for interference management
Medium access

- Interference management
  - Co-ordination of transmissions of nodes so that
    - interfering nodes are not transmitting simultaneously
    - and, overall wireless resource is utilized efficiently
  
  - In practice, this has to be achieved so that
    - each node makes decision using only local information
      - no access to ‘geographic clustering’ for tree
    - by means of simple algorithms

→ Need a medium access algorithm or protocol
Medium access

- Let’s play a game (of David McDonald)
  - Reward
    - 5 USD
Medium access

- Let’s play a game (of David McDonald)
  - Reward
    - 5 USD

- Rules
  - Respond, when asked, within 200ms
  - No reward if
    - none, or more than one simultaneous responses
  - Else, unique responder wins

- Fact: reaction time to auditory stimulus is 140-160ms
- Constraints
  
  - Interfering nodes cannot transmit simultaneously
• Constraints
  ○ Interfering nodes can not transmit simultaneously
Model

- Constraints
  - Interfering nodes cannot transmit simultaneously
  - Nodes have only local information
    - Contending simultaneous transmissions
Model

- Medium access
  - When to transmit subject to inference constraints
    - using local information
    - with an aim to maximize utilization of wireless medium
Model

- Network interference graph $G = (V, E)$ with $n$ queues
  - $E = \{(i, j) : i$ and $j$ can’t tx simultaneously\}$
  - Packets arrive at rate $\lambda_{id}$ at node $i$ for destination $d$
- Medium access: at each time instance
  - Selects non-interfering queues (to tx), i.e. independent set of $G$
Model

- Network interference graph $G = (V, E)$ with $n$ queues
  - $E = \{(i, j) : i$ and $j$ can't tx simultaneously\}$
  - Packets arrive at rate $\lambda_{id}$ at node $i$ for destination $d$
- Medium access: at each time instance
  - Selects non-interfering queues (to tx), i.e. independent set of $G$
Model

- Network interference graph $G = (V, E)$ with $n$ queues
  - $E = \{(i, j) : i$ and $j$ can't tx simultaneously\}$
  - Packets arrive at rate $\lambda_{id}$ at node $i$ for destination $d$
- Medium access: at each time instance
  - Selects non-interfering queues (to tx), i.e. independent set of $G$
Model

- Network interference graph $G = (V, E)$ with $n$ queues
  - $E = \{(i, j) : i \text{ and } j \text{ can't tx simultaneously}\}$
  - Packets arrive at rate $\lambda_{id}$ at node $i$ for destination $d$
- Medium access: at each time instance
  - Selects non-interfering queues (to tx), i.e. independent set of $G$
Model

- Let $\mathcal{I}(G)$ be set of independent sets of $G$
  - That is, $\mathcal{I}(G) = \{\sigma \in \{0, 1\}^n : \sigma_i + \sigma_j \leq 1 \text{ for all } (i, j) \in E\}$

- Effective service rate vector $\mu = [\mu_i]$ is s.t.
  - $\mu = \sum_{\sigma \in \mathcal{I}(G)} \alpha_{\sigma} \sigma$, with $\alpha_{\sigma} \geq 0$
    - $\sum_{\sigma} \alpha_{\sigma} \leq 1$

- Therefore, effective resource or ‘capacity region’
  - Convex hull of $\mathcal{I}(G)$, say $\text{conv}(\mathcal{I}(G))$
Performance metric

- Unicast demand $\lambda \in \mathbb{R}_{+}^{n \times n}$ is feasible if
  - There exists routing and medium access algorithm so that
    - induced per node demand, say $\hat{\lambda} \in \mathbb{R}_{+}^{n}$
    - is such that $\hat{\lambda} \in \text{conv}(\mathcal{I}(G))^{o}$

- Medium access algorithm is efficient if
  - Queues remain finite for any $\lambda$ that is feasible
    - formally, network Markov process is positive recurrent
Performance metric

- Notations
  - $Q(t) = [Q_{id}(t)] \in \mathbb{R}^{n \times n}$ be the queue-sizes at time $t$
  - $A(s, t) = [A_{id}(s, t)]$ cumulative arrivals to $i$ for $d$ in time $[s, t]$
    - arrival rate vector $\lambda = [\lambda_{id}]$
    - and $\mathbb{E}[A_{id}(s, t)] = \lambda_{id}(t - s)$
  - $\sigma(t) = [\sigma_{ij,d}(t)]$ be the schedule at time $t$
    - $\sigma_{ij,d}(t) = 1$: $i$ transmits packet for $d$ to $j$ at time $t$.
    - with $\hat{\sigma}(t) \in \mathcal{I}(G)$ where $\hat{\sigma}_i(t) = \sum_{j,d} \sigma_{ij,d}(t)$

- Dynamics: for each $i, d$

$$Q_{id}(t) = Q_{id}(s) + A_{id}(s, t) - \int_s^t \sum_j \left( \sigma_{ij,d}(y) \cdot 1_{Q_{id}(y) > 0} - \sigma_{ji,d}(y) \cdot 1_{Q_{jd}(y) > 0} \right) dy$$
Prior work

- Two classes of medium access algorithms (since early 70s)
  - Practical random access algorithm
  - Performance optimal queue-based algorithm
Random access algorithm

- Each queue $i$ checks medium ‘regularly’
  - Whether any ‘neighboring’ node is txing or not
  - If medium is free, attempts transmission with prob. $p_i$
    - upon being successful, tx for time duration $W_i$
  - Else
    - do nothing

- Popular back-off protocols are instance of this, e.g.
  - ALOHA [Abramson-Kuo 73], [Metcalf-Bogg 76]
Random access algorithm

- Properties
  - Naive message passing and easy to implement
  - But, as is poor in performance or hard to analyze

- Various positive, negative results
  - History (till 00) maintained by L. Goldberg
    http://www.csc.liv.ac.uk/~leslie/contention.html

- A notable, positive result by Hastad-Leighton-Rogoff 96
  - Polynomial backoff is throughput optimal
  - But, only for complete interference graph
Maximum weight algorithm

\[ t = 0 \]

\[ Q_3 = 45 \]
\[ \sigma_{31}, \sigma_{32} = ? \]
\[ Q_2 = 30 \]
\[ \sigma_{23}, \sigma_{25} = ? \]
\[ Q_5 = 0 \]

\[ Q_1 = 30 \]
\[ \sigma_{14}, \sigma_{13} = ? \]
\[ Q_4 = 10 \]
\[ \sigma_{41}, \sigma_{45} = ? \]

- Proposed by Tassiulas-Ephremides (1992)
  - For simplicity, assume only one destination, say 5
  - At time \( t \), choose \( \sigma(t) \) so that

\[
\sigma(t) = \arg \max_{\sigma \in \mathcal{I}(G)} \sum_{ij} \sigma_{ij} \left( Q_i(t) - Q_j(t) \right)
\]
Maximum weight algorithm

\[ t = 0 \]

\[ Q_3 = 45 \]
\[ \sigma_{31} = \sigma_{32} = 0 \]

\[ Q_2 = 30 \]
\[ \sigma_{23} = 0, \sigma_{25} = 1 \]

\[ Q_5 = 5 \]
\[ \sigma_{52} = \sigma_{54} = 0 \]

\[ Q_1 = 30 \]
\[ \sigma_{13} = 0, \sigma_{14} = 1 \]

\[ Q_4 = 10 \]
\[ \sigma_{41} = \sigma_{45} = 0 \]

- Proposed by Tassiulas-Ephremides (1992)
  - For simplicity, assume only one destination, say 5
  - At time \( t \), choose \( \sigma(t) \) so that

\[
\sigma(t) = \arg \max_{\sigma: \hat{\sigma} \in \mathcal{I}(G)} \sum_{ij} \sigma_{ij} \left( Q_i(t) - Q_j(t) \right)
\]
Generalized maximum weight algorithm

- At each time instance $t$
  - Choose $\sigma(t)$ so that
    $$\sigma(t) = \arg \max_{\sigma \in \mathcal{I}(G)} \sum_{ij} \sigma_{ij} \left( f(Q_i(t)) - f(Q_j(t)) \right)$$
  - For increasing function $f$ with $f(0) = 0$, $\lim_{x \to \infty} f(x) = \infty$

- Properties
  - Efficiency (positive recurrence) (Tassiulas-Ephremides (1992))
  - $f(x) = x^\alpha$ with $\alpha \to 0^+$ leads to smaller queues
    (Shah-Wischik (2006, 09))
  - But
    - as is, centralized and
    - requires solving computationally hard problem each time
• Our algorithm
  ○ Adaptive random access based on queue-size
  ○ ‘Simulates’ maximum weight algorithm
Our algorithm

- Each queue $i$ checks medium ‘regularly’
  - Whether any ‘neighboring’ node is txing or not
  - If medium is free, attempts transmission with prob. $p_i$
    - upon being successful, tx for time duration $W_i$
  - Else
    - do nothing

- Our choice: $p_i = 1$
  - tx to $j$ with min’l $Q$ and $\mathbb{E}[W_i] = \left(f(Q_i) - f(Q_j)\right)^+$
    - choice of $f$ determines performance crucially
    - a reasonable choice of $f$ is $\log$
Our algorithm

- Each queue has an independent Bernoulli clock of rate $1/2$

- If clock of queue $i$ ticks at time $t$, then
  - If $\sigma_{ij}(t-1) = 1$ for some $j$
    $$\sigma_{ij}(t) = \begin{cases} 0 & \text{with probability } \frac{1}{\left(f(Q_i(t))-f(Q_j(t))\right)^+} \\ 1 & \text{otherwise} \end{cases}$$
  - Else, $i$ check if medium free at time $t-1$
    - if so, attempt tx to $j$ with prob. 1 where $j$ is nbr of $i$ with minimal $Q$ and $Q_j(t) < Q_i(t)$
      $$\sigma_{ij}(t) = \begin{cases} 1 & \text{if no collision} \\ 0 & \text{otherwise} \end{cases}$$
Our algorithm: example

\[ Q_1 = 40 \]
\[ \sigma_1 = 1 \]

\[ Q_2 = 10 \quad \sigma_2 = 0 \]
\[ Q_3 = 5 \quad \sigma_3 = 0 \]
Our algorithm: example

$u.p. \frac{1}{2}$

$Q_1 = 40$
$\sigma_1 = 1$

$Q_2 = 10$  $\sigma_2 = 0$

$Q_3 = 5$  $\sigma_3 = 0$
Our algorithm: example

\[ \frac{1}{2} \]

\[ Q_1 = 40 \quad \sigma_1 = 1 \]

\[ Q_2 = 10 \quad \sigma_2 = 0 \quad Q_3 = 5 \quad \sigma_3 = 0 \]

\[ \frac{1}{f(40)} \]

\[ Q_1 = 39 \quad \sigma_1 = 0 \]

\[ Q_2 = 11 \quad \sigma_2 = 0 \quad Q_3 = 5 \quad \sigma_3 = 0 \]
Our algorithm: example

\[ Q_1 = 40 \]
\[ \sigma_1 = 1 \]

\[ Q_2 = 10 \quad \sigma_2 = 0 \]
\[ Q_3 = 5 \quad \sigma_3 = 0 \]

\[ Q_1 = 39 \]
\[ \sigma_1 = 0 \]

\[ Q_2 = 11 \quad \sigma_2 = 0 \]
\[ Q_3 = 5 \quad \sigma_3 = 0 \]
Our algorithm: example

\[ \begin{align*}
\text{w.p. } & \frac{1}{2} \\
\begin{align*}
\omega_1 &= 40 \\
\sigma_1 &= 1
\end{align*}
\end{align*} \]

\[ \begin{align*}
\omega_2 &= 10 \\
\sigma_2 &= 0
\end{align*} \]

\[ \begin{align*}
\omega_3 &= 5 \\
\sigma_3 &= 0
\end{align*} \]

w. p. \[ \frac{1}{f(40)} \]

\[ \begin{align*}
\omega_1 &= 39 \\
\sigma_1 &= 0
\end{align*} \]

\[ \begin{align*}
\omega_2 &= 11 \\
\sigma_2 &= 0
\end{align*} \]

\[ \begin{align*}
\omega_3 &= 5 \\
\sigma_3 &= 0
\end{align*} \]

w. p. \[ \frac{1}{4} \]

\[ \begin{align*}
\omega_2 &= 11 \\
\sigma_2 &= 0
\end{align*} \]

\[ \begin{align*}
\omega_3 &= 5 \\
\sigma_3 &= 0
\end{align*} \]

w. p. 1

\[ \begin{align*}
\omega_1 &= 39 \\
\sigma_1 &= 0
\end{align*} \]

\[ \begin{align*}
\omega_2 &= 11 \\
\sigma_2 &= 0
\end{align*} \]

\[ \begin{align*}
\omega_3 &= 6 \\
\sigma_3 &= 0
\end{align*} \]
Our algorithm: example

Q₁ = 39
Q₂ = 11
Q₃ = 5

Q₁ = 40
Q₂ = 10
Q₃ = 5

w.p. 1/2
w.p. 1/4
w.p. 1/3
Our algorithm: example

\[ Q_1 = 40 \quad \sigma_1 = 1 \]

\[ \text{w.p. } \frac{1}{2} \]

\[ Q_2 = 10 \quad \sigma_2 = 0 \]
\[ Q_3 = 5 \quad \sigma_3 = 0 \]

\[ \text{w.p. } \frac{1}{f(40)} \]

\[ Q_1 = 39 \quad \sigma_1 = 0 \]

\[ Q_2 = 11 \quad \sigma_2 = 0 \]
\[ Q_3 = 5 \quad \sigma_3 = 0 \]

\[ \text{w.p. } \frac{1}{4} \]

\[ \text{w.p. } 1 \]

\[ Q_2 = 11 \quad \sigma_2 = 1 \]
\[ Q_3 = 6 \quad \sigma_3 = 0 \]

\[ \text{w.p. } 1 \]

\[ Q_2 = 11 \quad \sigma_2 = 0 \]
\[ Q_3 = 6 \quad \sigma_3 = 0 \]

\[ \text{w.p. } \frac{1}{2} \]
Our algorithm: efficiency

• **Theorem.** [Ragagopalan-Shah-Shin 09, Shah-Shin 09, 10] The algorithm is efficient.
  
  o Use weight function \( f \) so that
    
    - \( f(x) = \exp(o(\log x)) \), like \( \log x \), \( \text{poly}(\log x) \), ... 

• Specifically, we establish that
  
  o The network Markov process is positive (Harris) recurrent
Best choice of $f$?

- Slower $f$ leads to
  - Small ‘variance’ in queue-sizes
  - At the cost of higher ‘average’ queue-sizes
    - price paid for network-wide ‘co-ordination’
Beyond capacity

- What about queue-sizes (on avg., with high prob.)?
  - For algorithm described, queue-sizes depend on
    - mixing time of random walk on space of schedules
    - could scale exponentially in number of nodes
  - But, for maximum weight schedule
    - Queue-sizes always scales polynomially in $n$

- Basic question: what are tradeoffs between
  - Capacity (or throughput), queue-sizes and complexity of algorithm
Beyond throughput

- Basic question: what are tradeoffs between
  - Throughput, queue-sizes and complexity of algorithm

- If algorithm is achieves at least 50% throughput, then
  - What is possible
    - Poly queue-size, but Exp complexity – maximum weight
    - Poly complexity, but Exp queue-size – our algorithm
  - What is *not* possible
    - Poly queue-size and poly complexity (Shah-Tse-Tsitsiklis (2009))
Beyond throughput

- If algorithm is achieves at least 50% throughput, then
  - What is possible
    - Poly queue-size, but Exp complexity – maximum weight
    - Poly complexity, but Exp queue-size – our algorithm
  - What is not possible
    - Poly queue-size and poly complexity (Shah-Tse-Tsitsiklis (2009))

- What about ‘regular’ networks
  - Random access for practical networks
    - with Poly queue-size?
  - Indeed, its possible (Shah-Shin (2010))
    - for network graphs with polynomial growth
    - at the expense of localized co-operation
Discussion

● We started with two questions
  ○ What is the network capacity?
  ○ How to build capacity achieving networks practically?

● Progress towards this questions
  ○ Scaling laws answer the first question for regular networks
  ○ Distributed, queue-based medium access provides
    ▪ means to achieve network-wide interference management
    ▪ for Gaussian fading with $\alpha > 3$ or Physical model
Going forward

- Two broad directions
  
  - Scaling laws for heterogeneous networks
    - with nodes with varied capabilities
    - and arbitrary locations
  
  - Distributed and efficient medium access to achieve network-wide
    - interference management and sophisticated co-ordination
    - e.g., to achieve capacity for Gaussian setup with $\alpha \in (2, 3)$
Related work

• An alternative approach [Jiang-Walrand 08, 09]
  ○ Given $\lambda \in \Lambda^o$
    ▪ find access probabilities $p(\lambda)$ s.t.
    ▪ resulting service rates $s(p(\lambda)) > \lambda$
  ○ MAX-ENT distribution satisfying these requirements
    ▪ exists and product-form
    ▪ parameters are appropriate dual variables
    ▪ yield to a sub-gradient algorithm

• Parameters can be learnt without any knowledge (rate stability)
  ○ Using appropriate incremental learning [Jiang-S-Shin-Walrand 09]
Related works

- Some of the recent related works
  - Modiano-S-Zussman 06
  - Gupta-Stolyar 06, Marbach 06
  - Duvry-Dousse-Thiran 07
  - Bordenave-McDonald-Proutiere 08
  - Liang-Walrand 08, Rajagopalan-S 08
  - Liang-Walrand 09, Liu-Yi-Proutiere-Chiang-Poor 09
  - Leconte-Ni-Srikant 09
  - Liang-S-Shin-Walrand 09
  - Liang-Walrand 10
  - S-Shin 10
  - van de Ven-van Leeuwaarden-Denteneer-Janssen 10
  - . . .