Self-Organizing Relays in LTE networks: Queuing analysis and algorithms

R. Combes[1] Z. Altman[1] E. Altman[2]

[1]Orange Labs

[2]INRIA Sophia-Antipolis

7th International Conference on Network and Service Management, 2011

Outline

Background and motivation

- Self-organization in communication networks has been recognized as a major axis of research by standardization bodies (e.g 3GPP) and researchers.
- A self-organizing network (SON) shall be able to perform configuration, optimization and troubleshooting tasks in an autonomic manner.
- Heterogeneous networks (HetNet) have been introduced in standards such as LTE-A. Low power nodes (pico-cells, femto-cells and relays) are deployed in heavy traffic areas to prevent outage.
- Because of the expected large number of small nodes, HetNets need SON functionalities
- In this work, we propose algorithms for self-organizing relays taking into account network traffic dynamics, i.e arrivals and departures of users.

Related work

Some related contributions are:

- The study of flow-level dynamics in wireless networks was studied in ([1])
- Scheduling in Multi-hop networks was discussed by the seminal ([2])
- The policy gradient approach to reinforcement learning was introduced in ([3])
- An extension to the average cost criteria was given in ([4])

[1] T. Bonald and A. Proutière. "Wireless Downlink Data Channels: User Performance and Cell Dimensioning". In: *ACM Mobicom*. 2003.

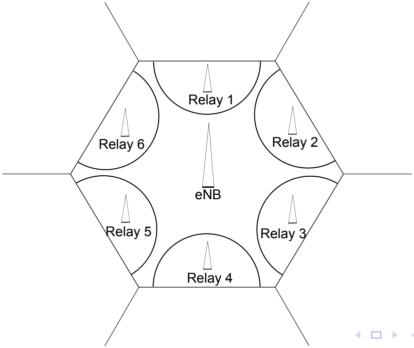
[2] L. Tassiulas and A. Ephremides. "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks". In: *IEEE Transactions on Automatic Control* (1992).

[3] Ronald J. Williams. "Simple statistical gradient-following algorithms for connectionist reinforcement learning". In: *Machine Learning* (1992).

[4] J. Baxter and P. L. Bartlett. "Infinite-Horizon Policy-Gradient Estimation". In: *Journal of Artificial Intelligence Research* (2001).

System model

- Downlink of a wireless network, serving elastic traffic (FTP-like)
- Users arrive at random instants and locations
- One base station (BS) and several relay stations (RSs). The BS to RSs links are wireless.
- $x\%$ of the radio resources are used for the BS to RSs links.



Flow-level Capacity

Each layer has a different definition of the capacity:

- PHY capacity
  - Single user, bandwidth W
  - Channel: AWGN
  - Capacity: $W \log_2(1 + SINR)$
- MAC capacity
  - N users sharing bandwidth W fairly
  - Channel: AWGN + fading, perfect CQI
  - Capacity: $\frac{1}{N}\mathbb{E}\left[W \log_2(1 + \overline{SINR}|Z|^2)\right]$
- IP (flow-level) capacity
  - Elastic (FTP-like) traffic
  - Call arrival rate λ, mean file size $\mathbb{E}[\sigma]$
  - BS s covers the zone $\mathbb{A}_s$
  - Capacity: $\left(\int_{\mathbb{A}_s} \frac{1}{R_s(r)} dr\right)^{-1}$

System capacity

Theorem

The capacity $C(x)$ of the system is

$$\min\left(x\left(\sum_{s=1}^{N_R} \frac{A_s}{R_{rel,s}}\right)^{-1}, (1-x)\max_{0\leq s\leq N_R}\left(\int_{\mathbb{A}_s} \frac{1}{R_s(r)}dr\right)^{-1}\right)$$

Furthermore, there exists a unique $x^* \in [0,1]$ which maximizes the capacity

$$x^* = \frac{\left(\max_{0\leq s\leq N_R} \int_{\mathbb{A}_s} \frac{1}{R_s(r)}dr\right)^{-1}}{\left(\max_{0\leq s\leq N_R} \int_{\mathbb{A}_s} \frac{1}{R_s(r)}dr\right)^{-1} + \left(\sum_{s=1}^{N_R} \frac{A_s}{R_{rel,s}}\right)^{-1}}$$

Relaying Gain

- We assume that relays are placed high enough so that the BS to RS signal experiences line-of sight attenuation $\frac{A}{\|r\|^2}$, while the interfering signals attenuation are $\frac{A}{\|r\|^{\eta r}}$. We call this difference in attenuation relaying gain.
- We evaluate the capacity gain, with and without relaying gain. It is shown that without relaying gain, relays do not improve the network capacity.

Model parameters	
Cell layout	Hexagonal
Antenna type	Omnidirectional
Cell Radius	$2km$
Access technology	OFDMA
Fast-fading model	Rayleigh
N_{RB}	10
Resource block size	$180kHz$
BS transmit power	$46dBm$
RS maximum transmit power	$30dBm$
Thermal noise	$-174dBm/Hz$
Path loss model	$128 + 37.6\log_{10}(d)$ dB, d in km
File size	$10Mbytes$

Table: Model parameters

Numerical experiments

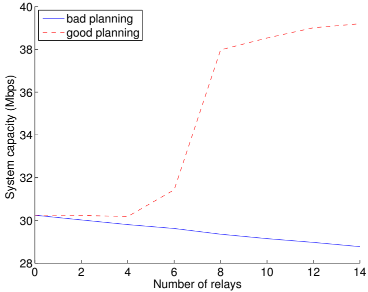


Figure: System capacity for different planning strategies

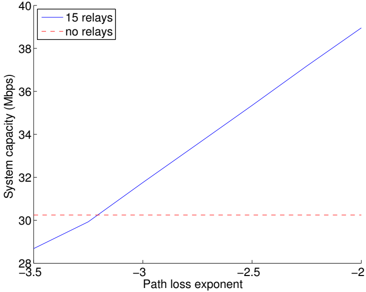


Figure: Impact of the relaying gain on the system capacity

Markov decision processes(MDPs)

MDP's model the problem of taking sequential decisions in an uncertain environment.

- State space $\mathcal{S}$, Action space $\mathcal{A}$ (both discrete)
- Reward $r : \mathcal{S} \to \mathbb{R}$
- Transition probabilities: $\mathbb{P}\left[s_{t+1} = s | s_t, a_t\right]$
- Policy $P : \mathcal{S} \to \mathcal{A}$, $a_t = P(s_t)$
- To each policy P we associate a cost:
  - (average cost) $J_{\mathbf{s}_0}(P) = \lim_{T \to +\infty} \mathbb{E}_{P,\mathbf{s}_0}\left[\frac{1}{T}\sum_{t=0}^{T} r_t\right]$
  - (discounted cost) $K_{\mathbf{s}_0,\gamma}(P) = (1-\gamma)\mathbb{E}_{P,\mathbf{s}_0}\left[\sum_{t\in\mathbb{N}} \gamma^t r_t\right]$
- Aim: find the policy that minimizes the cost
- Optimal policy can be found using value iteration if $|\mathcal{S}|$ and $|\mathcal{A}|$ are small.

Formulation as a MDP

- The BS can observes the buffer size of each link, and chooses which links to activate
- MDP model:
  - State space: $\mathcal{S}$ buffer size each link
  - Action space: $\mathcal{A} = \{0, 1\}$ activate user links/backhaul
  - Cost: mean user delay, using Little's law: $\lim_{T \to +\infty} \left[\frac{1}{T} \sum_{t=0}^{T} n_t \right]$
- The MDP is tractable for a small number of RSs

Scalable resolution methods for large MDPs

- For large state spaces, the problem becomes non-tractable ("curse of dimensionality")
- Solution: introduce a "well-chosen" subset of policies:
  - Based on the optimal policy for a small version of the problem
  - Allows generalization and can be used for reinforcement learning
- Policy parameter $\theta \in \mathbb{R}^N$, P_θ associated policy, cost $J(\theta) = J_{\mathbf{s}_0}(P_\theta)$
- Cost $J(\theta)$ is minimized using classical optimization techniques

Parametrized Policies

- We introduce a set of parametrized *deterministic* policies:

$$P_{d,\theta}(\mathbf{S}, 1) = \begin{cases} 1 & , \quad < \mathbf{S}, \theta > \ \geq 0 \\ 0 & , \quad < \mathbf{S}, \theta > \ < 0 \end{cases} \tag{1}$$

- And a set of parametrized *stochastic* policies:

$$P_{s,\theta}(\mathbf{S}, 0) = 1 - f(< \mathbf{S}, \theta >)\,, \text{ with } f(x) = \frac{1}{1 - e^{-x}} \tag{2}$$

Numerical experiments

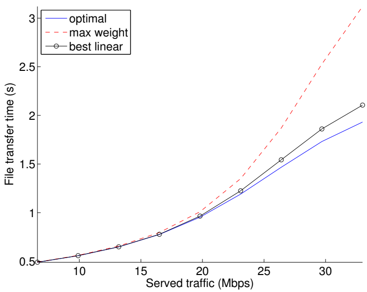


Figure: File transfer time as a function of the traffic for different control strategies

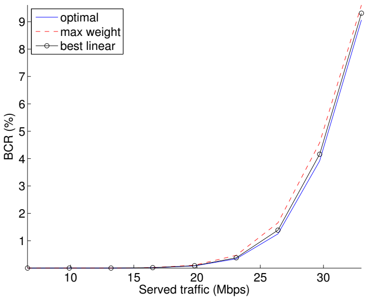


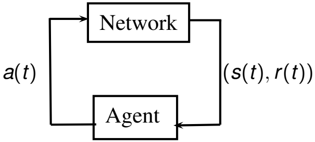
Figure: Block call rate as a function of the traffic for different control strategies

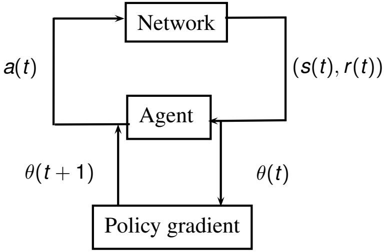
Reinforcement learning

- A MDP is considered, but transition probabilities are unknown, “model-free learning”
- An agent observes states, actions and rewards to make a decision $\{s_t, a_t, r_t\}_{t \in \mathbb{N}}$.
- Basic approach: Q-learning, but impractical for large MDPs

Policy gradient approach

- Used approach: policy gradient, we use the previous family of policies
- $\nabla_\theta J(\theta)$ can be estimated from observations (Baxter and Bartlett).
- Policy parameter θ is updated using gradient descent: $\theta_{n+1} = \theta_n + \epsilon_n \nabla_\theta J(\theta_n)$





Implementation issues

- The learning method is valid regardless of the statistical assumptions on traffic.
- Since the policy gradient applies to MDP's and partially observed MDP's.
- Scalable approach: components of $\nabla_\theta J(\theta_n)$ are estimated from the same observations $\{s_t, a_t, r_t\}_{t\in\mathbb{N}}$.
- Scalability is key since some deployment scenarios include 30 RSs per BS.

Numerical experiments

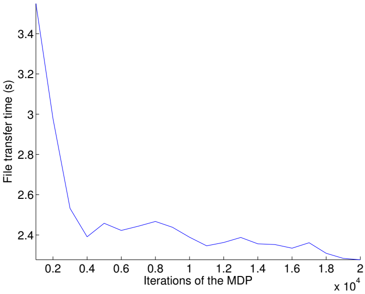


Figure: File transfer time during the learning process

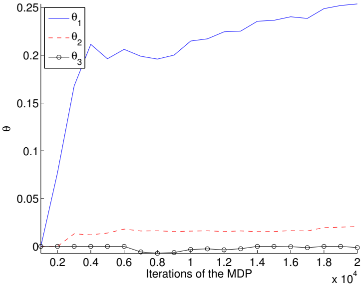


Figure: Controller parameters $(\theta_1, \theta_2, \theta_3)$ during the learning process

Conclusion

- We have considered the problem of self-organized relays in a cellular network
- The capacity of the system has been derived using queuing theory
- Dynamic resource sharing has been considered to optimize blocking rate and file transfer time
- A parametrized set of policy has been introduced, to maintain scalability
- A model-free approach has been shown, in which the network can derive an optimal parametrized policy, based on observation and interaction with the network