A study on machine learning and regression based models for performance estimation of LTE HetNets

B. Bojović[1], E. Meshkova[2], N. Baldo[1],
J. Riihijärvi[2] and M. Petrova[2]

[1]Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Castelldefels, Spain

[2]Institute for Networked Systems, RWTH Aachen University,
Aachen, Germany

ACROPOLIS 3rd Annual Workshop - London, 2013

Outline

1. Motivation
   - Femtocell network deployments
   - Inter-cell interference awareness
   - Dynamic frequency allocation for LTE system
2. Problem statement
   - LTE physical layer
   - Proposed prediction methods and covariates
3. Simulation scenario
4. Results of performance prediction
   - Issues in predicting performance
   - Machine learning techniques for performance estimation

Femtocell network deployments

Goals and issues.

- Recently, there was much interest in femtocells deployments for different network technologies.
- The most of attention was in shown for WCDMA technology in 3G network deployments
- The same concept could be applied also to more recent technologies such as WiMAX, LTE,..
- The goal is to improve the performance of the mobile network by:
  - extending indoor coverage
  - achieving higher data rate for the indoor and short range outdoor communications
- The main challenge is how to manage the available spectrum resources

Femtocell network deployments.

Goals and issues.

- One approach could be to divide the available spectrum into several frequency bands and then to assign to each femtocell a different one
- In previous case, the drawback are: low or no frequency reuse, not feasible for dense deployments and requires much maintenance
- Another approach, which is typically considered in recent technologies, is to share available frequency bands among femtocells
- In the previous case the main issue is the interference between femtocells, between femtocell and macro cells
- Consequence: the network performance degradation and unfairness among users

Inter-cell interference awareness

Previous work

- Many studies have identified significant performance gains in the systems that are aware of the inter-cell interference
- Objective: to maximize the system performance and to provide fairness among users by minimizing the inter-cell interference
- Some of proposed techniques are based on: power control schemes, static and adaptive fractional frequency reuse schemes, cancellation techniques, intelligent scheduling, network power coordination,...

Dynamic frequency allocation for LTE system
Our approach

- Our approach it to improve the system performance by enabling the dynamic frequency allocation
- We consider an LTE system, which is designed to be used with frequency reuse factor of 1
- LTE technology includes advanced features such as OFDMA, SC-FDMA, AMC, dynamic MAC scheduling and HARQ,..
- More difficult than in previous technologies to predict the system capacity for specified system configuration
- Our approach is to apply machine learning and regression analysis for system capacity estimation, that will enable efficient dynamic frequency allocation

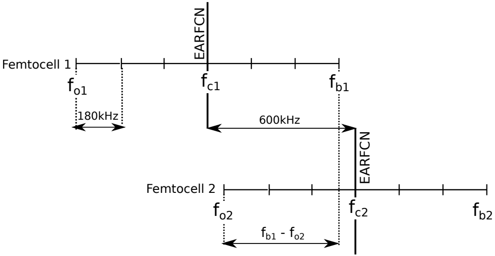
LTE physical layer
Configurable network parameters

- The OFDMA multicarrier technology provides flexible multiple-access scheme that allows different spectrum bandwidths to be utilized without changing the fundamental system parameters or equipment design
- In this work we consider:
  1. the carrier frequency f_c; and
  2. he bandwidth B
- In frequency domain resources are grouped in units of 12 subcarriers that are called resource blocks centered around f_c and occupy 180 kHz in frequency domain

Configuring carrier frequency and bandwidth.

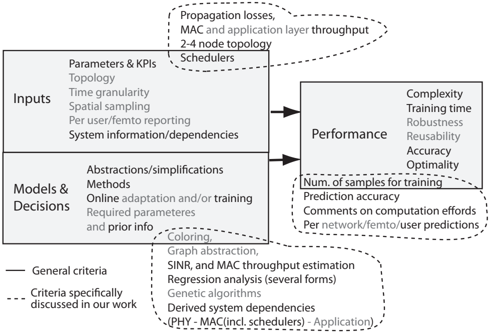
An example

- The channel raster is 100 kHz for all bands
- LTE eNB operates using a set of B contiguous resource blocks(RB); the allowed values for B are 6, 15, 25, 50, 75, 100 RBs (1.4 , 3, 5, 10, 15, 20 MHz)



The optimization problem in broader context

- The quality of the achieved performance depends on various factors: amount of available inputs, their usability in the context of particular models and optimization methods



The specific optimization problem

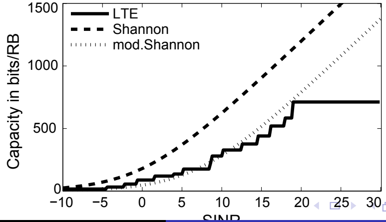
An example

- The specific problem that we are solving is: select for each for each deployed eNB $i = 1, ..., N$, the frequency f_c and the system bandwidth B that will provide the best network performance.
- In this work we consider capacity as performance metric (we are interested to consider delay, fairness, different QoS metrics)
- Another aspect to consider is the number of possible solutions, which is exponential with N
- An example for 4 femtocell scenario, total available bandwidth of 5 MHz, f_c muliple of 300 kHz and B that can take value 6,15,25 there are 4625 physically distinct solutions

LTE femtocell capacity estimation

design the cost function

- Common approach is to use some variation of Shannon formula, but those approaches typically do not consider effects of AMC, MAC scheduling,..
- The TB size for each given AMC scheme and number of RBs is defined by the LTE specification
- We consider effect of different schedulers and for that purpose we use Round Robin and Proportional Fair



prediction methods and covariates

- Our goal is to predict the performance of the femtocell accurately by using regression analysis and machine learning techniques
- The information used: basic pathloss and configuration information, a limited number of feedback measurements that provide the throughput and the delay metrics for a particular frequency settings
- Different regressors: SINR, SINR/MAC throughput mapping and different sampling technique affects the prediction performance
- The performance metrics considered are: network-wide and per-user throughput
- The methods considered: Bagging tree, Boosted tree, Kohonen network, SVM radial, K-nearest neighbor,

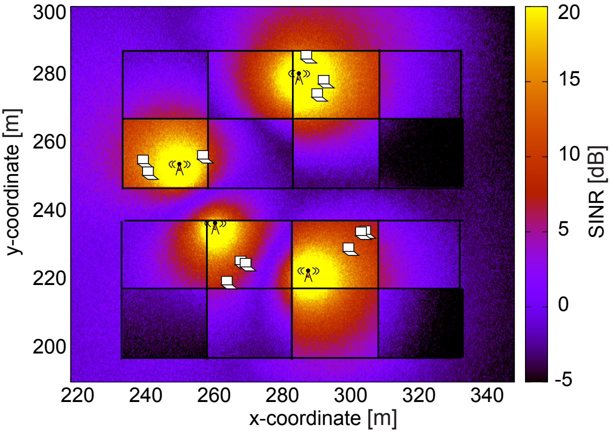
Simulation scenario.

- To simulate this scenario we use LENA, LTE-EPC network simulator
- We choose a typical LTE urban scenario with buildings (in LTE literature known as dual stripe scenario)
- The HeNbs are randomly distributed in the buildings and each HeNb has equal number of users
- For the four femtocell scenarios we simulate:
  1. three users per node allocation with the total of 2 MHz bandwidth, and
  2. two users per node allocation with the total of 5 MHz bandwidth

Centre Tecnològic de Telecomunicacions de Catalunya

Dual-stripe scenario

Radio environmental map in a scenario with 2 buildings and 2 HeNbs



Issues in predicting performance

Scenario 1

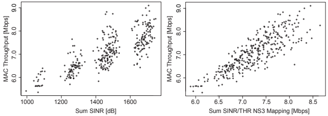
- Scenario with 4 femtocells and 3 users each femto, using 2 MHz overall bandwidth, showing actual measured MAC throughput vs. Shannon models. Sum SINR means a sum over all RBs
- Even if correlations between the different SINR related metrics and the achieved throughputs are apparent, the dispersion of the results is rather substantial

CTTC Centre Tecnològic de Telecomunicacions de Catalunya

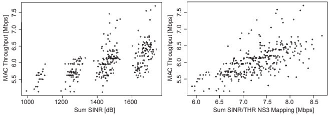
RWTH AACHEN UNIVERSITY

Issues in predicting performance

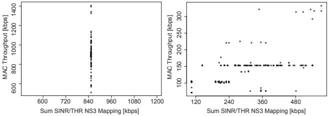
Scenario 1



(a) Measured MAC throughput vs. min SINR (left) and SINR/MAC thoughput mapping (right) for a scenario with 4 femtos with 3 users each, Proportional Fair scheduler, and TCP traffic.



(b) Measured MAC throughput vs. min SINR (left) and SINR/MAC thoughput mapping (right) for a scenario with 4 femtos with 3 users each, Round Robin scheduler, and TCP traffic.

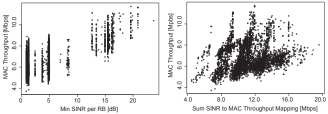


(c) Measured MAC thoughput for two selected users vs. SINR/MAC THR mapping for a scenario with 4 femtos 3 users each, Proportional Fair (left) and Round Robin schedulers (right) for TCP traffic.

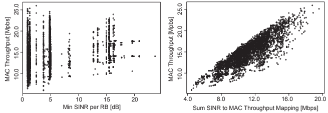
Issues in predicting performance
Scenario 2

- Scenarios with 2 femtocells and 5 users each and 4 femtocells with 2 users each, using a bandwidth of 5 MHz.
- Different behavior is noted for UDP and TCP traffic, and for different schedulers
- From these results we can expect effective performance prediction to be a challenging problem especially for TCP traffic and a small number of users per femtocell.

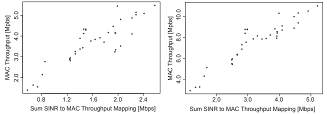
Issues in predicting performance
Scenario 2



(a) Measured MAC throughput vs. min SINR (left) and SINR/MAC thoughput mapping (right) for a scenario with 4 femtos with 2 users each, Proportional Fair scheduler, and TCP traffic.



(b) Measured MAC throughput vs. min SINR (left) and SINR/MAC thoughput mapping (right) for a scenario with 4 femtos with 2 users each, Proportional Fair scheduler, and UDP traffic.



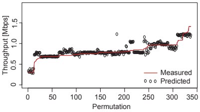
(c) Measured MAC throughput vs. SINR/MAC thoughput mapping for a selected femtocell (left) and a whole network (right) for a scenario with 2 femtos with 10 users each, Round Robin scheduler, and UDP traffic.

Pursuit regression technique

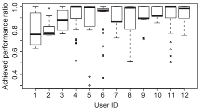
- Scenario with 4 femtocell and 3 users per each femto and bandwidth of 2MHz
- Random sampling with 10% of 337 permutations being explored
- The regression technique is applied
- The earlier expectation that UDP traffic with more primitive scheduler results in higher predictability is confirmed, the users the predictions are very accurate

Pursuit regression technique

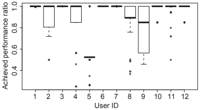
Figure



(a) Measured and predicted MAC throughput for a selected user with permutations ordered after measured throughput for the scenario with TCP traffic, and proportional fair scheduler.



(b) Achieved performance ratio (predicted vs. measured MAC throughput) for the scenario with TCP traffic, and proportional fair scheduler.

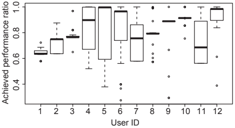


(c) Achieved performance ratio (predicted vs. measured MAC throughput) for the scenario with UDP traffic, and round robin scheduler.

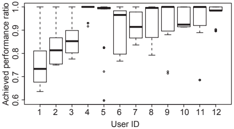
A simple linear regression model

- Scenario with 4 femtocell scenario and 3 users per femto and bandwidth of 2 MHz
- TCP traffic and proportional fair scheduler
- This method results in poorer prediction average both in terms of median error, as well as the variability of the results
- The increased amount of information available for the predictor results in both improved median prediction performance, as well as substantial reduction in the magnitude of the outliers

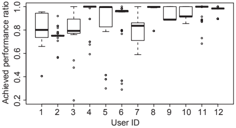
A simple linear regression model
Figure



(c) Achieved performance ratio for the linear regression method with 10% of samples and the stratified sampler.



(c) Achieved performance ratio for the Kohonen regression method with 10% of samples and the random sampler.



(c) Achieved performance ratio for the PPR regression method with 70% of samples and the stratified sampler.

Conclusions

- The initial results obtained here show that advanced regression techniques can result in good performance predictions
- In future work we plan to investigate an active sampling based on graph coloring