Faculty of Engineering, Built Environment and Information Technology

Fakulteit Ingenieurswese, Bou-omgewing en Inligtingtegnologie / Lefapha la Boetšenere, Tikologo ya Kago le Theknolotši ya Tshedimošo

Machine Learning-based Base Station Association for Resource Allocation in 5G Heterogeneous Cognitive Networks

M.M. Sande, M.C. Hlophe, and B.T. Maharaj

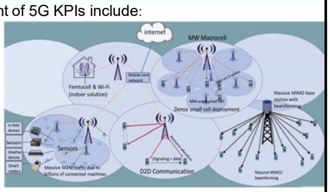
31 May 2018

Presentation Outline

- Overview of 5G networks
- Modelling for devices in 5G networks
- Machine learning
- System model & problem formulation
- Q-learning-based BS association for CRN
- Simulation results

5G Networks Overview

- 5G aims to support connection of the internet of people and internet of things, while providing high data rates and reduced latency for communication
- Technologies that enable achievement of 5G KPIs include:
  - Cognitive radio (CR) terminals/networks
  - Massive MIMO & mmWave
  - Heterogeneous networks (HetNets)
  - Device-to-device (D2D) communication
  - Energy harvesting
  - Smart grids



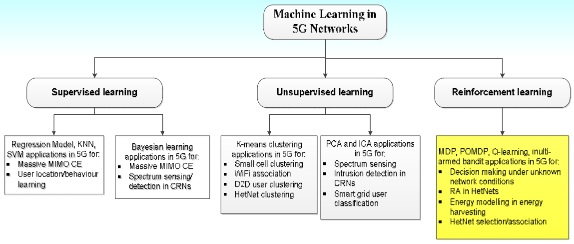
Modelling for Optimizing 5G Networks

- Smart devices for 5G networks are capable of jointly using many different frequency bands and communication technologies.
- Analysis of network characteristics such as user association, resource allocation and data routing using SINR as the base metric for resource allocation is indequate.
- There is need for integrating multiple parameters in the optimization of 5G networks.
- Consider context-aware models for cognitive user terminals, which enable interference management and avoidance in multi-tier 5G networks

Machine Learning (ML)

- A modelling technique that has found applications in
  - Gaming
  - Data mining
  - Telecommunications
  - Bio-sciences
  - Control automation
- An artificial intelligence tool that is being conceived to support smart CR terminals
- Basically explained, a machine learns the execution of a particular task with the goal of maintaining a specific performance metric, based on a particular experience.

Machine Learning in 5G



System Model

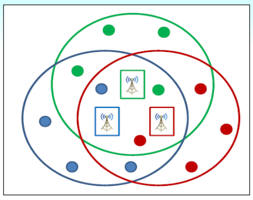


Fig. 1. System model showing an overlay deployment of three small cell clusters in a macrocell

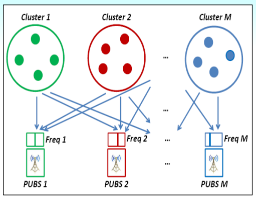


Fig. 2. Illustration of clustering in BS association

Problem Formulation

The optimization problem for resource allocation is formulated as

$$U = \max \sum_{i \in N} x_{i,m} \cdot f^{(m)}, \qquad (1)$$

subject to

$$\textstyle\sum_{i \in N} \lambda_s^i \leq C, \qquad (2)$$

$$\psi_{i,m} \geq \psi_{min}, \qquad \forall i \in N, \qquad (3)$$

where $x_{i,m} = \left(\frac{b_m}{\sum_{i \in N} \lambda_s^i}\right) \log(1 + \psi_{i,m})$ is the throughput for user i, provided by BS m,

C is the capacity of each PUBS (no. of SU transmissions that it can accommodate)

ψ_{min} is the minimum QoS requirement for all SUs

$f^{(m)}$ is a classification rule that maps previous BS load and cluster states to a future network state

Q-learning

- A reinforcement learning technique that does not require the exact transition formulation of the system
- Maintain a Q-table consisting of Q values that represent a reward resulting from taking an action a when in state s
- An agent is trained such that its actions should interact with the environment to maximize the cumulative reward resulting from the interactions

9

Q-learning Based BS Association in CRN

- The objective in this scenario is to minimize the infinite horizon discounted power cost, subject to a constraint on the infinite horizon discounted delay
- The expected discounted power cost and expected discounted delay are then defined as

$$\bar{P}^{\pi}(s) = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i \rho\left(s^i, \pi(s^i)\right) | s^0 = s\right] \text{ and}$$

$$\bar{D}^{\pi}(s) = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i g\left(s^i, \pi(s^i)\right) | s^0 = s\right],$$

respectively. The formal objective is then stated as

$$\text{minimize } \bar{P}^{\pi}(s),$$

$$\text{subject to } \bar{D}^{\pi}(s) \le \delta,$$

where δ is the discounted buffer cost constraint.

10

Q-learning Algorithm for a CRN Context

Algorithm 1: Proposed Approximate Q-Learning Algorithm for Base Station Association in a CRN

Input: x_0, θ, α
Output: Trained SUs
01: Initialize x_0 and θ
02: **For** each learning episode **do**
03: Arbitrarily initialize $Q(s, a)$
04: Initialize s_t
05: **For** each step t **do**
06: Choose a from $\mathcal{A}$ using derived ϵ-greedy
07: Take action a_t, observe r, s_{t+1}
08: $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$
09: Measure x_{t+1}, receive r_{t+1}
10: $\theta \leftarrow \theta + \alpha_t[r_{t+1} + \max_{\alpha_{t+1}} Q(x_a^{t+1}, a_{t+1}; \theta) - Q(x_t, a_t; \theta)] \frac{\partial \hat{Q}(x_t, a_t; \theta)}{\partial \theta}$
11: **End for**
12: $s_t \leftarrow s_{t+1}$
13: **End for**

11

Simulation Results

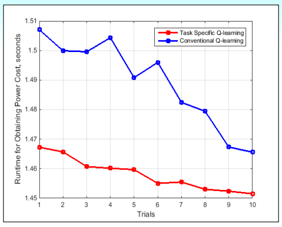


Fig. 3. Comparison of runtime for gathering power cost for the proposed task specific Q-learning and conventional Q-learning

Parameter	Value
Bandwidth	8 MHz
Downlink center frequency	2.1 GHz
Channel model type	Jakes
Scheduling algorithm	Proportional fair
Inter-site distance	20 m
Propagation environment	Urban
BS maximum transmit power	46 dBm
Shadow fading margin standard dev	8 dB
Noise power density	-174 dBm/Hz
Neural network hidden layer dimension	24, 24
One episode duration	50 TTIs
Discount factor, γ	0.950
Learning rate, α	0.001

12

Conclusion

- A machine learning-based user association policy for SUs in a cluster-based CRN was presented.
- The proposed task specific Q-learning technique gave better runtime convergence compared to the conventional Q-learning technique in obtaining the cost of transmitting at a certain power.
- The proposed approach is more favorable for application in wireless networks where there are not enough computational resources to smartly learn more than one aspect at every time slot.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Thank You

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA