2024.05.09

Cell On/Off Parameter Optimization for Saving Energy via Reinforcement Learning

Mobile Communication Lab Weekly Seminar

Index Terms:
Cell on/off, energy-saving, network, artificial intelligence

2023310331 Geon Kim

MCL
KYUNGHEE

KYUNG HEE
UNIVERSITY

OUTLINE

1. Why this paper?
2. Summary of paper
3. Take away

MCL
KYUNGHEE

01

Why this paper?

Contributions

- **Formulating an RL problem controlling a cell on/off algorithm in a way to minimize energy consumption while satisfying throughput constraints**

- Proposing a range of operational modes on top of the trained RL agent

- Presenting experimental results with a replicative simulator

01

Summary of paper

MCL KYUNGHEE

Introduction

- The energy that base stations consume has increased compared to the legacy system, thus placing a high burden on service providers in **OPEX (OPeration Expendigure)**

- Along with the OPEX reduction, saving energy decrease **carbon emission**, thus, preserve environments

- This paper address the energy saving of base stations by **turning off cells**, while **not severely impacting network performance**

Reinforcement Learning

- Turning the cells off varies depending on the traffic, channel characteristics, and mobility pattern of the users associated with the cells

- Designing a universal algorithm that can be applied to all circumstances is almost infeasible. Also, each network operator has its own criteria on service level

- The energy saving needs to be **personalized** and **adaptive** to the circumstanceds as well as the network operator’s requirements.

- It is infeasible to try out unproven controls on actual RAN equipment for it would degrade user experience by interrupting the service
  - Most reinforcement learning based problem formulations make use of simulations

- Author utilizes a replicative simulator that has a capability to reconstruct the states and the behaviors of an actual RAN based on the observed data from it

- To match with the LTE configuration of the target region, **four spectrum bands** each with the corresponding spectrum characteristics are virtually implemented

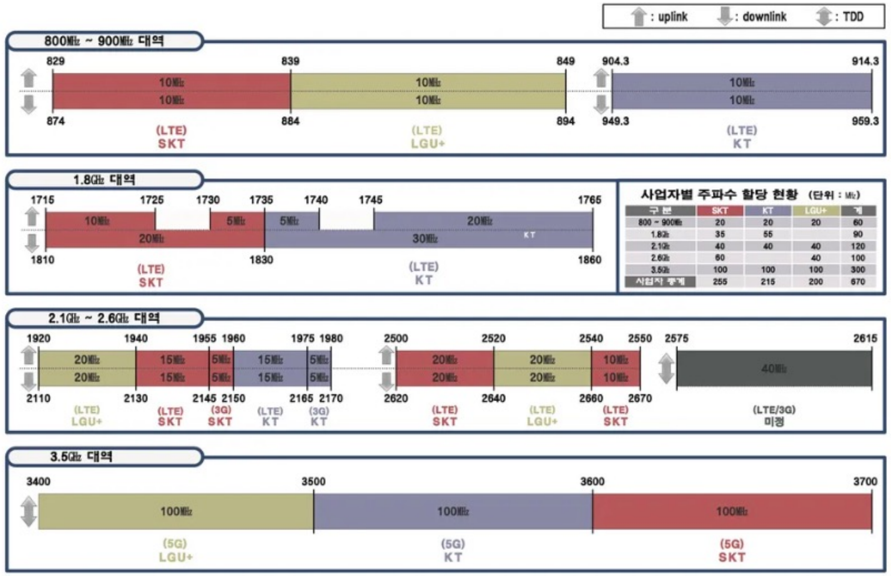

Figure 1. Frequency band allocation for South Korean telcos

System Model – Cell On/Off Algorithm

- Cell on/off algorithm which proposed in this paper works with a couple of **threshnold values on the PRB utilization**
    1. Activation threshold
    2. Deactivation threshold

- After the cell activation or deactivation, UEs are distributed among the activated cells by a load balancing algorithm

- It is crucial to find **the appropriate threshold values** for the algorithm to operate optimally depending on the state of the network

System Model – Power Model

- The RAN topology consists of three sectors on a single site and four cells per sector (4 * 3 = 12 cells)
  - Each cell is served by a separate **RU with a different power consumption profile** depending on serving bandwidth and hardware settings
- Power consumption linearly increases with the number of PRB in use
- Many factors to consider when deciding optimal thresholds for energy saving algorithm
  - Bandwidth, power model, hardware configurations, traffic pattern etc.

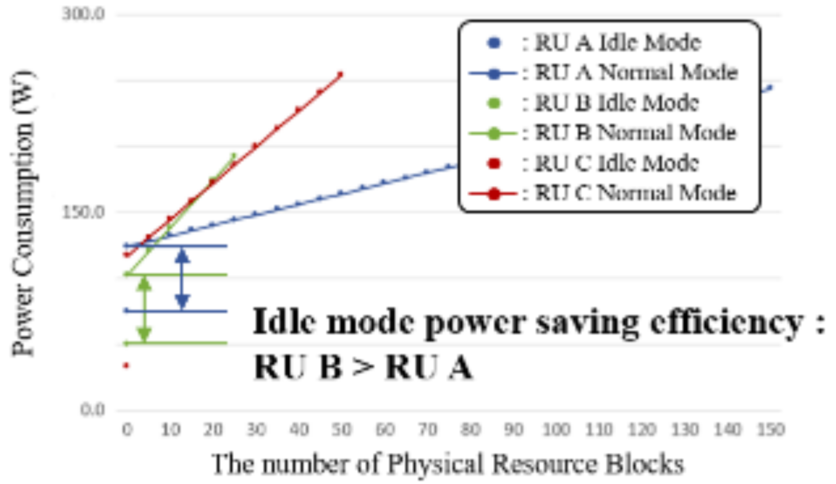

Figure 2. Power consumption trend according to active RBs per cell

Reinforcement Learning-Based Customized Energy Saving Policy - RL Model Design

- PPO (Proximal Policy Optimization), or DDPG (Deep Deterministic Policy Gradient) are used to derive an **optimal policy,** $\boldsymbol{\pi(a|s)}$

- Detailed formulation
  - **State**

$$s_t^i = (l_{t-1}^i, \dots, l_{t-K}^i, c_{t-1}^i, \dots, c_{t-K}^i)$$

    - l_t^i: Load of cell i at time k
    - c_t^i: On/Off state of cell i

  - **Action:**
    - Activation and deactivation threshold for cell i at time t

- Detailed formulation
  - **Reward**
    - Power reward: r_{power}

$r_{power} = \beta_0^p + \beta_1^p l_t^i$c

- β_0^p: 1
- β_1^p: $-1/(P_{max} - P_{min})$

- Throughput reward: r_{tput}

$$r_{tput} = \begin{cases} \beta_0^{tp} + \beta_1^{tp} k_t^i, & if\ k_t^i > Q \\ \beta_0^d + \beta_1^d k_t^i, & otherwise \end{cases}$$

- q: Minimum throughput constraint
- k_t^i: Throughput of cell i at time t
- β_0^{tp}: $-q/(T_{max} - q)$
- β_1^{tp}: $1/(T_{max} - q)$
- β_0^d: $-10q$
- β_1^d: 10√√

  - **Total reward**

$r_t = \sum(\alpha r_{power} + (1-\alpha) r_{tput})$

Where α is a real number satisfying $0 < \alpha < 1$

PPO (Proximal Policy Optimization)

- PPO is known to provide **robust performance** on the convergence of the training and support both the discrete and continuous action space

- Make use of M multiple simulation instances in a cluster of CPU machines in parallel to **generate a tremendous amount of required experiences**

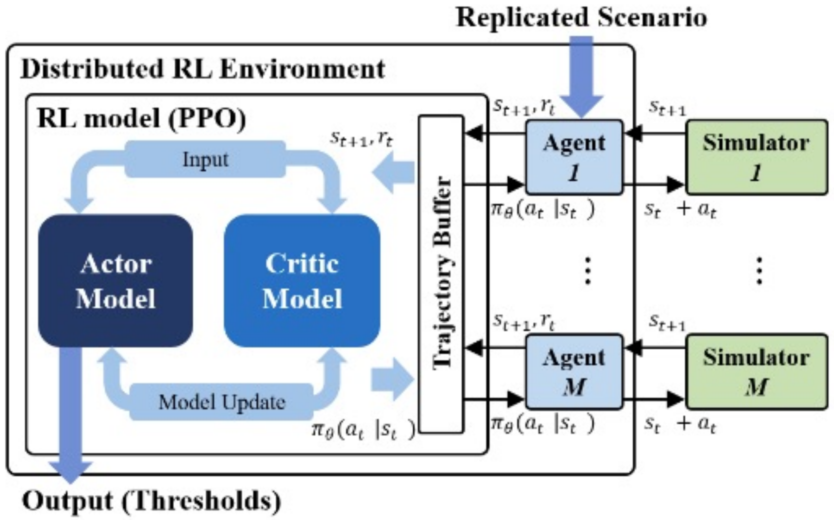

Figure 3. Reinforcement Learning based Energy Saving Model Architecture with Distributed Comoputing Techniques

Reinforcement Learning-Based Customized Energy Saving Policy – Various Operational Modes

- To balance between the energy saving performance and cost of running the AI-based solution, this paper explores various operational modes

- One extreme mode is to dynamically apply inferred actions from the trained models
  - Computation power at the backside of the system ↑
  - Requiring more resources

- Propose to operate the energy saving solutions whith scheduling tables of the thresholds generated offline
  - Tables can be created using trained models and a **network status prediction module** to forecast future network conditions

- By **varying periods for controls and prediction**, we can balance the **performance** and the **required computation power**

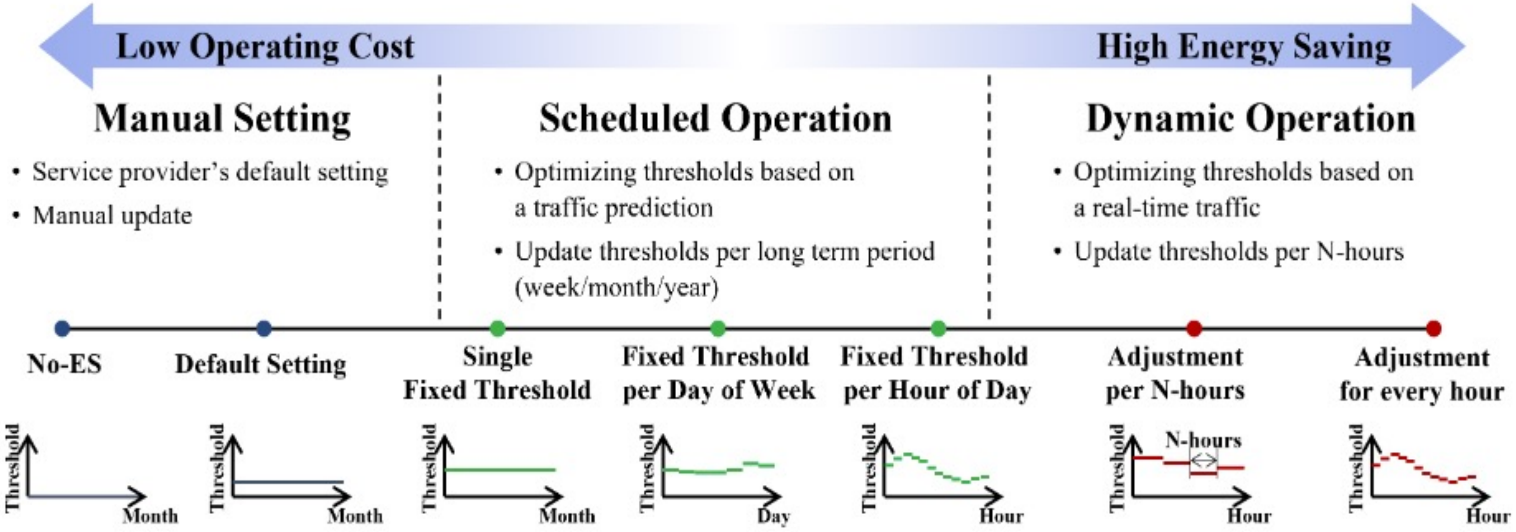

Figure 4. Range of operationg options

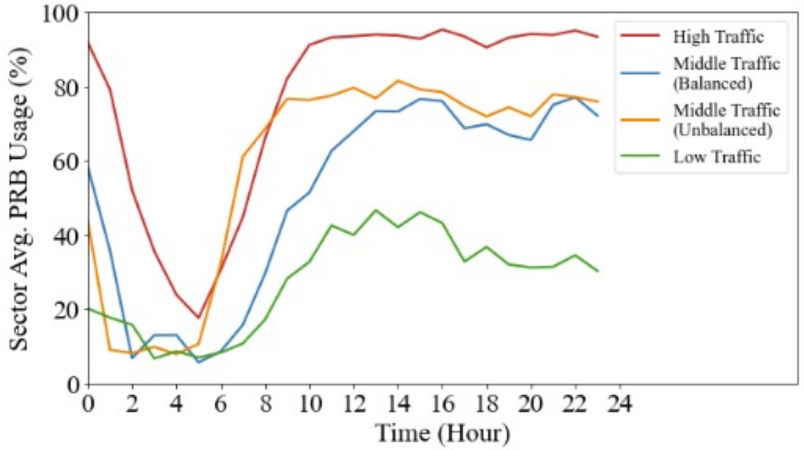

Figure 5. Traffic data used in the experiments

- Dynamic operation:
  - Maximizes energy-saving performance

- Offline-based operation
  - Cost-effective energy savings

Reinforcement Learning-Based Customized Energy Saving Policy – Evaluation Scenarios

- No-ES
  - All cells are activated for all time
- Aggressive/Conservative Daily Fixed Threshold
  - Aggressive algorithm
    - Higher values on both thresholds
    - Deactivate and activate: 60%, 80%
  - Conservative algorithm
    - Lower values on both thresholds
    - Deactivate and activate: 20%, 40%
- Sector/Cell-wise Daily Fixed Threshold
  - Thresholds are determined with the history of traffic load
  - Deactivate and activate: 25%, 75%
- RL-based Daily Fixed Threshold
  - Infers a single set of thresholds for a day
- RL-based Hourly Fixed Threshold
  - Infers threshold values per hour for a scenario

Reinforcement Learning-Based Customized Energy Saving Policy – Evaluation Results

Scenario	Algorithm	Power	Throughput	Violation
Low Traffic	No-ES	449 W	7.65 Mbps	0.00%
	Conservative	−1.86%	−8.66%	+0.00%p
	Aggressive	−4.89%	−57.01%	+9.03%p
	Sector-wise	−0.52%	−1.49%	+0.00%p
	Cell-wise	−0.71%	−1.94%	+0.00%p
	RL Daily	−11.87%	−11.94%	+0.00%p
	RL Hourly	−11.90%	−8.21%	+0.00%p
Middle Traffic (Balanced)	No-ES	513 W	2.59 Mbps	0.35%
	Conservative	−1.48%	−0.77%	+0.34%p
	Aggressive	−2.48%	−7.34%	+2.43%p
	Sector-wise	−1.83%	+0.39%	−0.35%p
	Cell-wise	−1.07%	+3.09%	−0.35%p
	RL Daily	−5.68%	−6.18%	−0.35%p
	RL Hourly	−5.99%	−9.65%	+0.34%p
Middle Traffic (Imbalanced)	No-ES	535 W	1.84 Mbps	3.82%
	Conservative	−0.74%	−2.17%	+0.35%p
	Aggressive	−1.38%	−4.35%	+1.04%p
	Sector-wise	−0.52%	−1.09%	−1.04%p
	Cell-wise	−0.92%	−3.26%	+1.04%p
	RL Daily	−2.36%	−3.26%	+0.69%p
	RL Hourly	−3.48%	−6.52%	+0.00%p
High Traffic	No-ES	602 W	0.59 Mbps	30.21%
	Conservative	+0.44%	+1.69%	+2.78%p
	Aggressive	−0.61%	+0.00%	+1.73%p
	Sector-wise	+0.33%	+0.00%	+1.39%p
	Cell-wise	+0.43%	+0.00%	+1.39%p
	RL Daily	−1.65%	+0.00%	−1.74%p
	RL Hourly	−2.45%	+1.69%	−2.43%p

- 1 MBPS minimum throughput constraint
- Violation" means the time portion in which the constraints are not satisfied

- Energy saving performance depends on the scenario's **traffic pattern**, such as traffic volume and traffic distribution among cells
  - After a cell is turned off, other activated cells need to serve users not to incur throughput violations
  - **Network status needs to be considered holistically** when deciding to activate or deactivate cells, indicating that our RL-based operations are required

Algorithm	Power	Violation (2 Mbps)
No-ES	535 W	18.4 Mbps
RL Hourly (1 Mbps)	−3.48%	+3.13%p
RL Hourly (2 Mbps)	−3.02%	+0.00%p

- 1 Mbps vs. 2 Mbps minimum throughput constraint

Conclusion

- Problem
  - Energy saving problem: **Minimizing energy comsumption** while **guaranteeing a given level of throughput** by controlling the parameters of a heuristically designed algorithm

- Optimization Framework : **Reinforcement learning**

- Network's behaviors and states are **replicated with a simulator** that reconstructs real world scenarios with the data collected from the equipment

- Evaluation results show that solution can achieve **maximum energy saving** while fulfilling the throughput requirement

01

Take away

MCL KYUNGHEE

Take away

1. 주파수 대역을 고려한 cell on/off 알고리즘은 이미 있음

2. 강화학습을 위해 필요한 파라미터들을 기억할 것
   - Cell의 load(traffic으로 대체 가능할까?)
   - Activation and deactivation threshold를 action으로 사용

3. Engineering에서 사용하는 강화학습은 실제 상황에 직접적으로 dynamic하게 적용하는 것이 아니라 offline으로 prediction한 데이터로 학습 후 적용시킨다는 점

Date: 2024, May, 10

Thank You