The University of Texas at Austin
Electrical and Computer Engineering
Cockrell School of Engineering

Wireless Networking & Communications Group

Q-Learning Algorithm for VoLTE Closed-Loop Power Control in Indoor Small Cells

Faris B. Mismar and Brian L. Evans

faris.mismar@utexas.edu and bevans@ece.utexas.edu

MOTIVATION — I

Downlink VoLTE Power Control (PC)

- Network impairments impact SINR and data rates
- LTE and 5G NR do not have downlink power control
- Link adaptation is instead used in the downlink
- Network intelligence enables clearer communications
- *Current practice*:
  - ❖ Diversity and more robust modulation and coding
- *Proposed*:
  - ❖ Reinforcement learning is used to maximize SINR

Goal

- Enable downlink power control for indoor voice bearer
- Data is 60% likely to be on Wi-Fi!

Approach

- Q-learning improves the downlink SINR through:
  - ❖ Performing exploration and exploitation
  - ❖ Running until it finds a near-optimal policy

NETWORK MODEL — II

- An indoor OFDM-based cellular cluster in the FDD mode of operation
- Multi-user multiple access with one tier of neighboring cells
  - ❖ Square geometry and inter-site distance of length L
- Network impairments regularly occurring at random
- PC keeps track of forward link SINR and potential radio impairments

Forward link budget for the i-th user

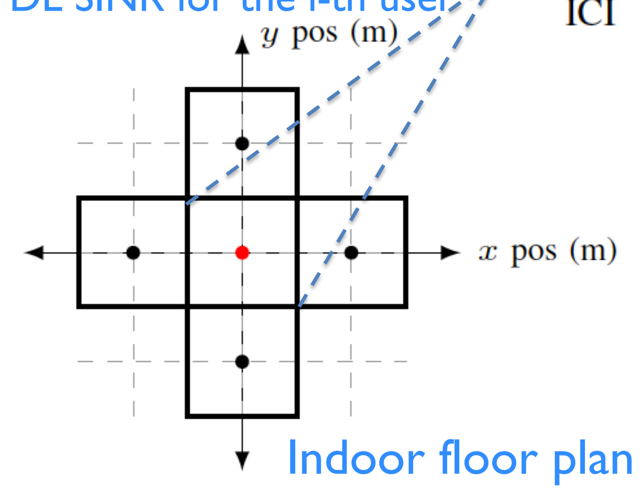
$$P_{\mathrm{UE},i}[t] = P_{\mathrm{TX}}[t] + G_{\mathrm{TX}} - L_{\mathrm{m}} - L_{\mathrm{a},i}[t] + G_{\mathrm{UE}}$$

$$\gamma_{\mathrm{DL},i} \triangleq \frac{|y_i|^2}{\sigma_n^2 + \underbrace{\sum_{j:\mathbf{o}_j \in \mathcal{C}\setminus\{\mathbf{o}_0\}} k_j |y_j|^2}_{\text{ICI}}}$$

DL SINR for the i-th user

$$P_{\mathrm{TX}}[t] = \min(P_{\mathrm{BS}}^{\max}, P_{\mathrm{TX}}[t-1] + \eta[t]c[t])$$

Power control repetition (up to 3 times per TTI)



Indoor floor plan

POWER CONTROL (PC) ALGORITHM ACTIONS AT TIME t

Action a	Definition
0	Nothing (this is a transient action).
1	Three (PC $= -1$) executed (i.e., $\eta[t] = 3$).
2	Single (PC $= -1$) executed (i.e., $\eta[t] = 1$).
3	Single (PC $= +1$) executed.
4	Three (PC $= +1$) executed.

Algorithm:
VoLTE Power Control

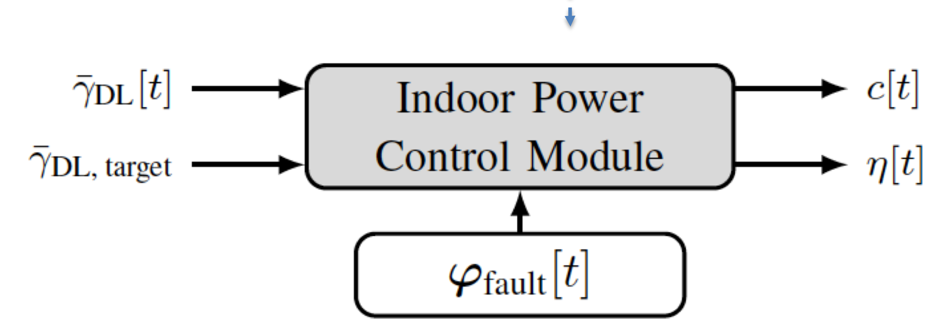
Given the set of fault-handling actions $a \in \mathcal{A}$ in a packetized voice frame z of duration τ.

1) Play an action a at random with ϵ probability, or based on max future reward with probability $1-\epsilon$
2) Change the downlink transmit power per a
3) Obtain the reward $r_{s,a}$ and the next state s'
4) Update s and $Q(s,a)$ accordingly
5) Decay ϵ and repeat for the next frame

PARAMETERS — III

Radio Network

Radio Network Parameters			
Cell radius	10 m	*Geometry*	Square
f_c	2600 MHz	*Antenna pattern*	Omnidirectional
Bandwidth	20 MHz	*Prop. model*	COST231
Tx Power	33 dBm	*Tx Ant. Height*	10 m
Tx, Rx branches	2,2	*Rx Ant. Height*	1.5 m



The reward

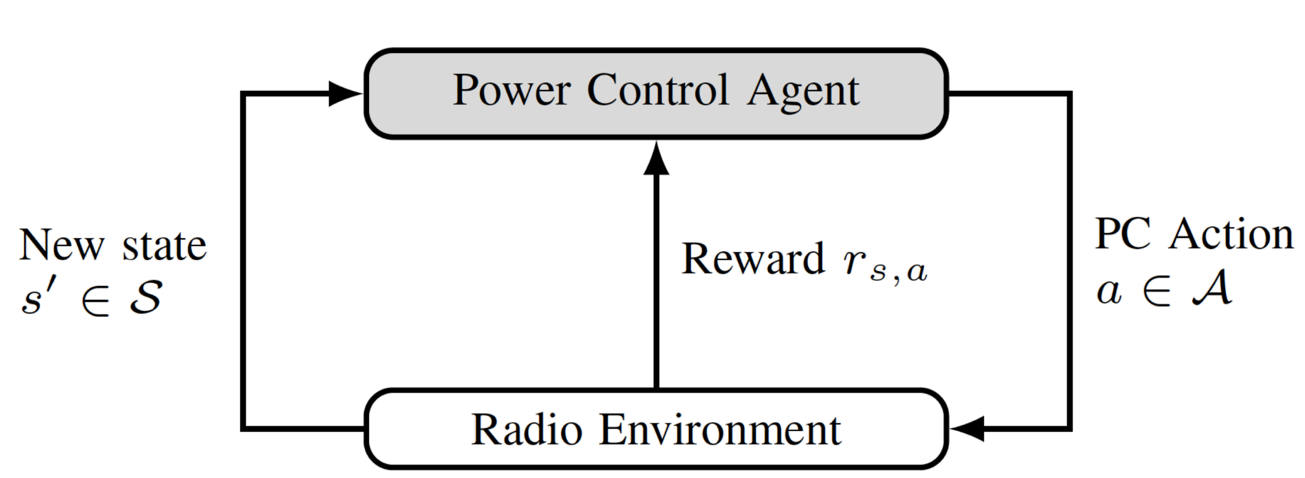
$$Q(s,a) \triangleq (1-\alpha)Q(s,a) + \alpha\left[r_{s,a} + \gamma \max_{a'} Q(s',a')\right]$$

The discounted future-looking reward

Q-Learning Parameters

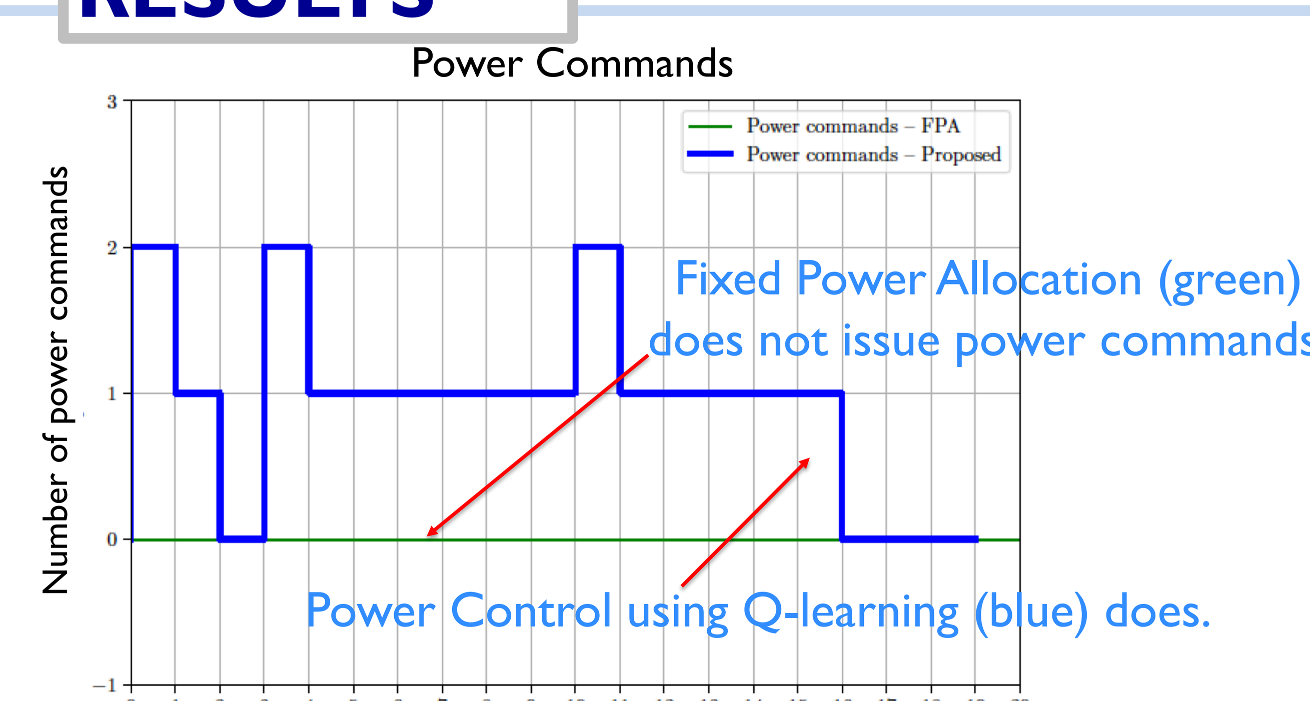
MACHINE LEARNING PARAMETERS

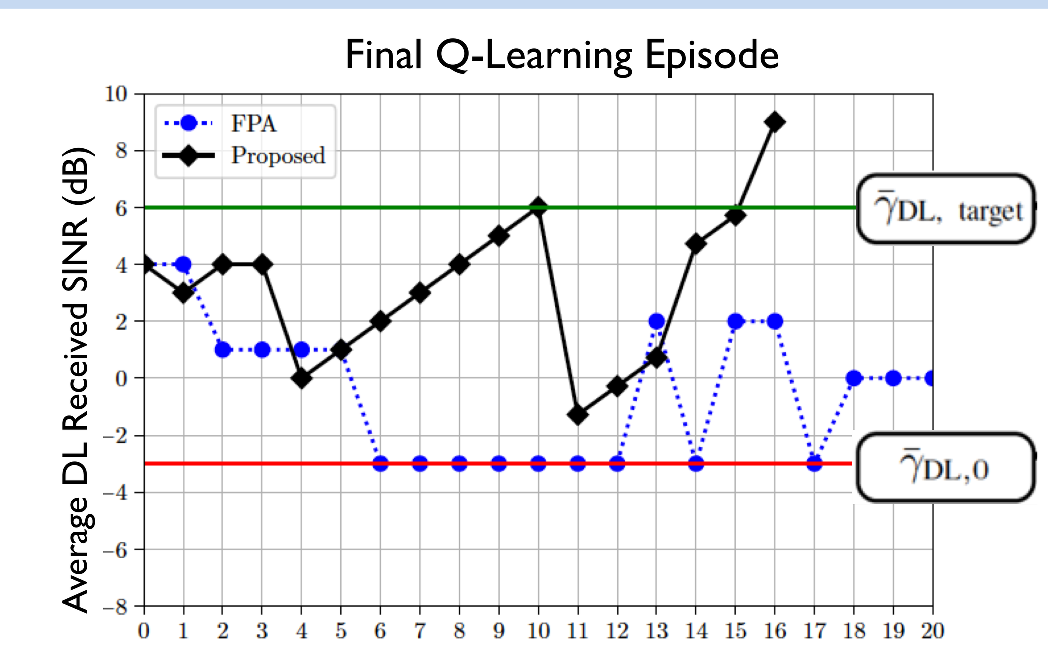
Parameter	Value
Number of episodes ζ	707
One episode duration τ (ms)	20
Discount factor γ	0.950
Exploration rate ϵ	1.000
Minimum exploration rate $\epsilon_{\min}$	0.010
Exploration rate decay d	0.99
Learning rate α	0.001
Number of states	3
Number of actions	5

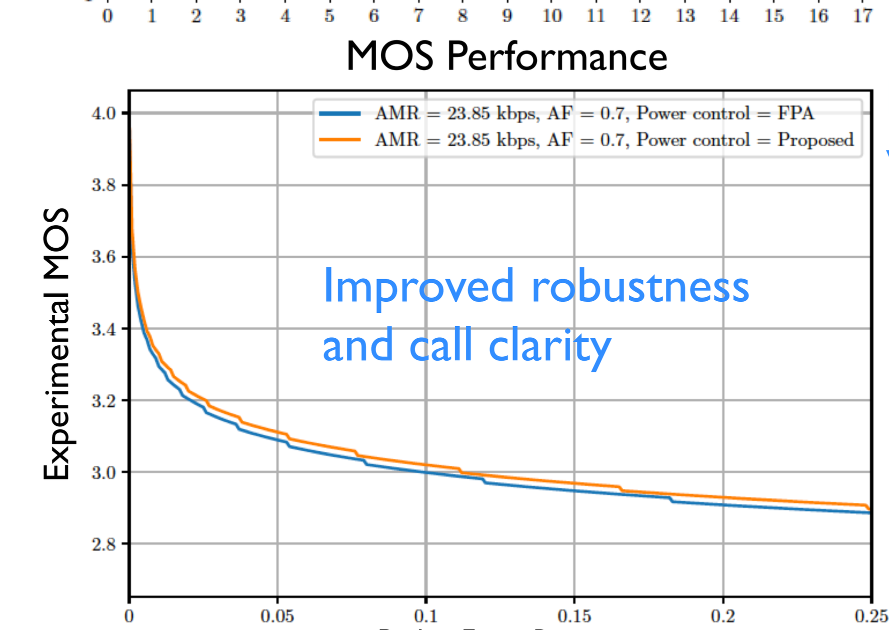


Reinforcement learning elements

RESULTS — IV







Retainability decreases when a call is dropped.

	Fixed Power Allocation	Proposed
Retainability	55.00%	**78.75%**

CONCLUSIONS — V

- Improved downlink SINR
- Achieved higher voice call retainability and clarity
- Reinforcement learning enables power control
- Compatible with LTE and 5G NR standards

October, 2018

Personal homepage
https://www.linkedin.com/in/farismismar/