The University of Texas at Austin
Electrical and Computer Engineering
Cockrell School of Engineering
Wireless Networking & Communications Group

Deep Q-Learning for Self-Organizing Networks Fault Management and Radio Performance Improvement

Faris B. Mismar and Brian L. Evans
faris.mismar@utexas.edu and bevans@ece.utexas.edu

ESPL

MOTIVATION — I

Self-Organizing Networks

- Cellular network faults impact SINR and data rates
- *Current practice*: manual intervention
  - ❖ RF engineer monitors network for faults
  - ❖ Manually attempts to clear alarms
- *Proposed*: self-healing self-optimizing network (SON)
  - ❖ Deep Q-network (DQN)
  - ❖ Learn near-optimal fault-handling sequence
- Network intelligence enables 5G rates

Goal

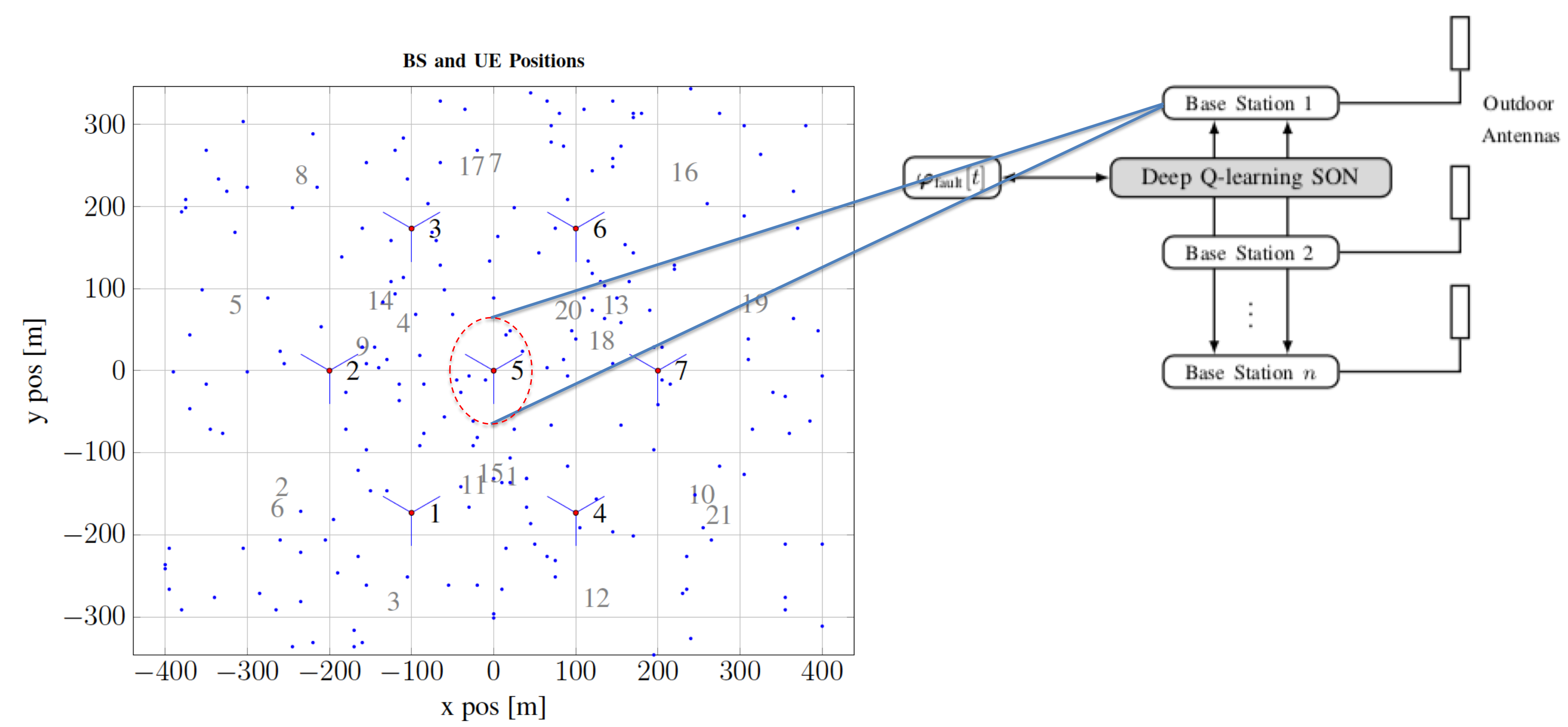
- Enable self-healing functionality in a network

Approach

- DQN improves downlink SINR through:
  - ❖ Performing exploration and exploitation
  - ❖ Running until it finds a near-optimal policy.

NETWORK MODEL — III

- Outdoor OFDM-based cellular cluster in FDD mode of operation
  - ❖Multi-user multiple access with one tier of neighboring cells
  - ❖Hexagonal geometry and inter-site distance of length L
- Network faults $\mathcal{N}$ regularly occur with no known optimal fault-handling policy
- Self-organization network (SON) tracks faults using a vector $\varphi_{\text{fault}}(t) \in \{0,1\}^{|\mathcal{N}|}$.



Algorithm: SON Fault Management

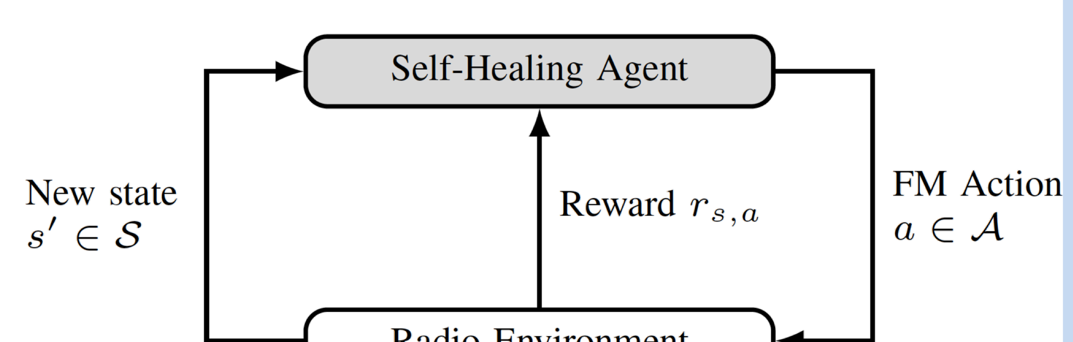
Given set of fault-handling actions $a \in \mathcal{A}$ in an LTE frame z of duration τ

1) Play action a at random with ϵ probability, or replay from buffer with $1-\epsilon$ probability
2) Update fault register if alarm is cleared by a
3) Obtain reward $r_{s,a}$, and next state s'
4) Store (s, a, r, s') in buffer and sample from it
5) Set $y_t \triangleq \mathbb{E}_{s'}\left[r_{s,a} + \gamma \max_{a'} Q(s', a'; \boldsymbol{\theta}_{t-1}) \,\middle|\, s, a\right]$
6) Do stochastic gradient descent on $(y_t - Q(s, a; \boldsymbol{\theta}_t))^2$
7) Update state s and decay ϵ
8) Repeat until terminal or exceed frame duration

PARAMETERS — II

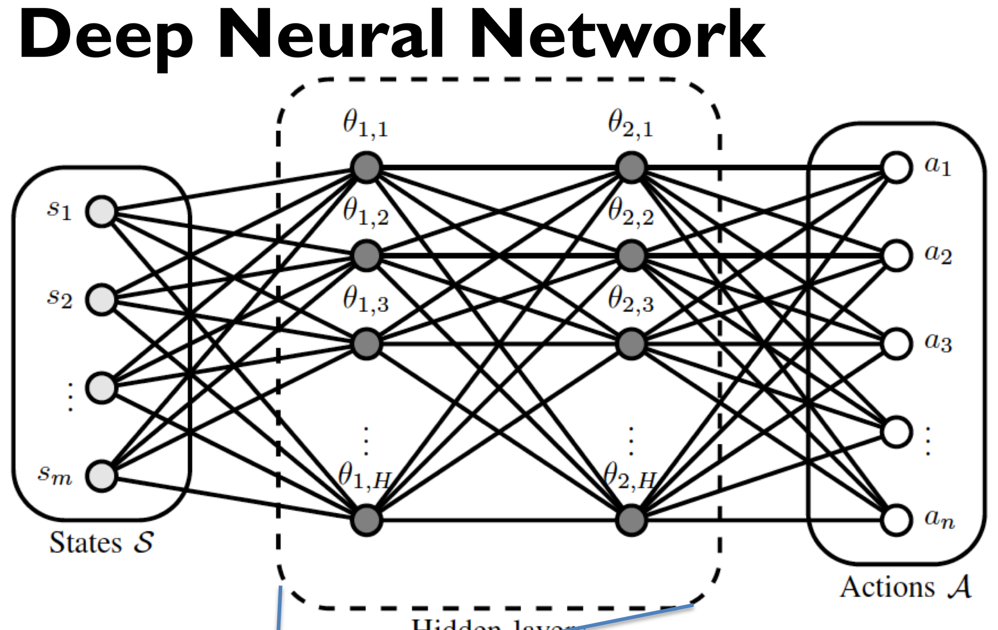
Radio Network

Radio Network Parameters			
Cell radius f_c *Bandwidth* *Tx Power*	200 m 2100 MHz 20 MHz 46 dBm	*Geometry* *Antenna pattern* *Prop. model* *Tx Ant. Height* *Rx Ant. Height*	Hexagonal 3gpp 36.942 COST231 20 m 1.5 m



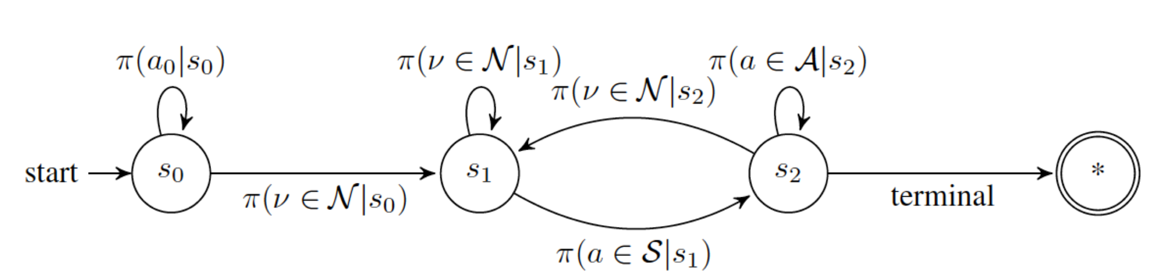
Reinforcement Learning Elements

Deep Neural Network



$Q(s, a; \boldsymbol{\theta}_t) \approx Q^*(s, a)$ as $t \to \infty$

The weights from the hidden layers



Markov Decision Process (MDP)

Obtained reward — The discount factor

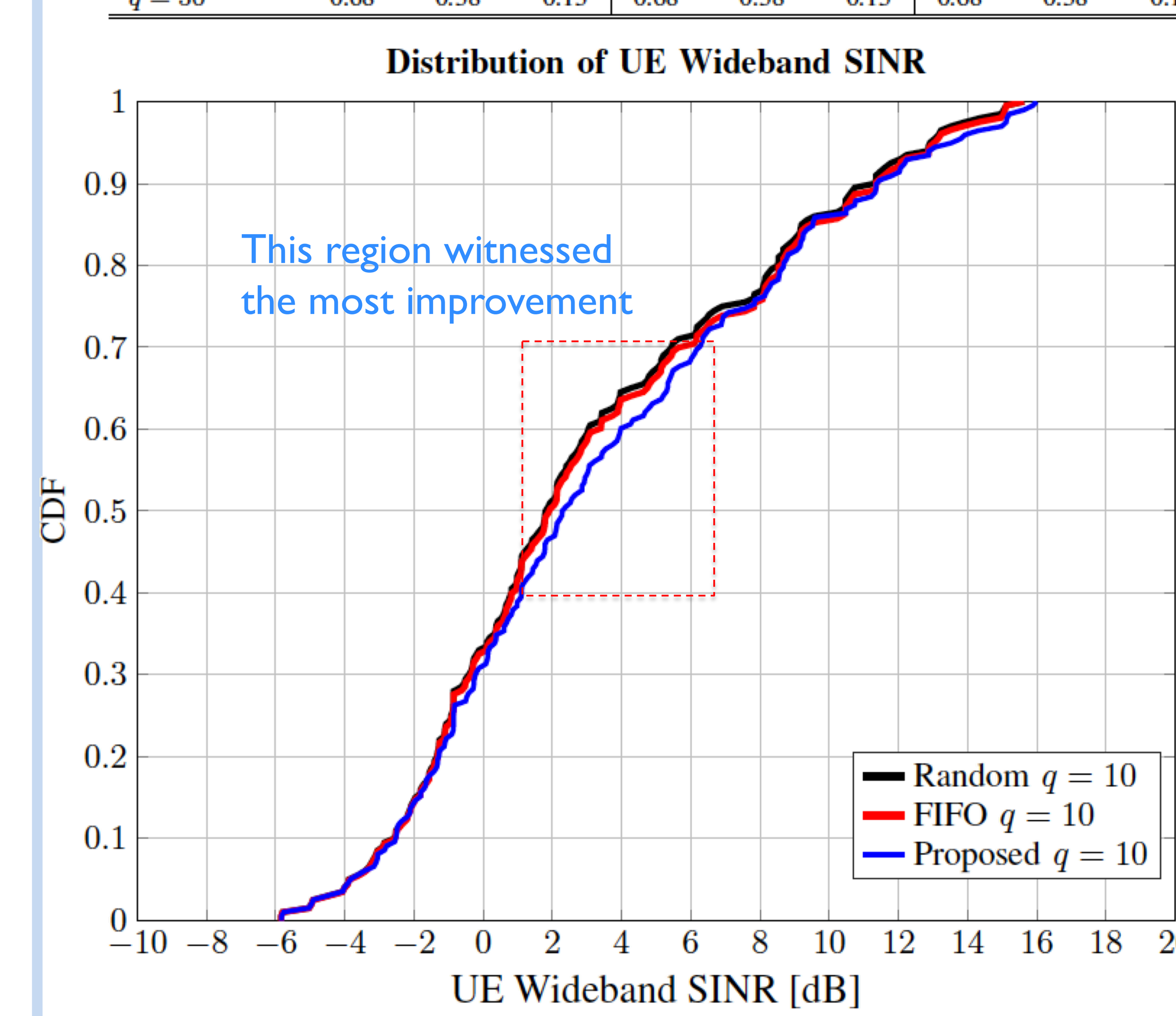
$$Q^*(s, a) \triangleq \mathbb{E}_{s'}\left[r_{s,a} + \gamma \max_{a'} Q^*(s', a') \,\middle|\, s, a\right]$$

The expected discounted reward (Bellman).

RESULTS — IV

THROUGHPUT PERFORMANCE PER ALGORITHM FOR DIFFERENT NUMBER OF USERS PER CELL

	User Equipment (UE) Throughput [Mbps]									Cell Throughput [Mbps]		
	Random			FIFO			Proposed			Random	FIFO	Proposed
UEs per cell	Peak	Average	Edge	Peak	Average	Edge	Peak	Average	Edge	Average		
$q = 10$	3.48	1.78	0.53	3.52	1.79	0.54	**3.55**	**1.84**	**0.58**	17.77	17.95	**18.37**
$q = 50$	0.68	0.38	0.13	0.68	0.38	0.13	0.68	0.38	0.13	18.81	18.96	**18.97**



The proposed algorithm wins for moderate number of users (q = 10).

Performance is comparable for large number of users (q = 50) due to resource depletion.

CONCLUSIONS — V

- Improved downlink SINR distribution
  - ❖ Leads to higher UE downlink rates.
  - ❖ Most efficient with moderate q.
- Self-healing capability in cellular networks:
  - ❖ Can be enabled through DQN.

October, 2018

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