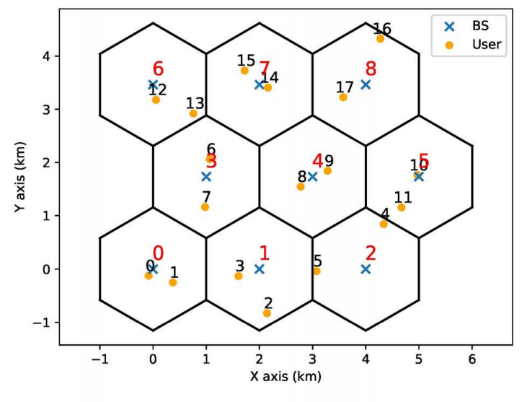
power allocation in multi-user cellular networks with deep q learning approach

cited by 23 times, [ICC 2019 - 2019 IEEE International Conference on Communications (ICC)](https://ieeexplore.ieee.org/xpl/conhome/8753818/proceeding)

1. Environment

Distributed dynamic downlink power allocation with multiple users and an interfering multiple-access channel(IMAC)

Contribution

1. Transfer learning(off-line train -> on-line train)

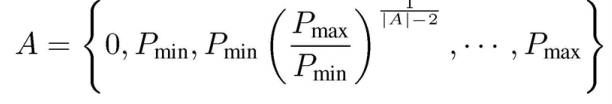
2. No future rewards

3. After centralized training, tested by distributed execution

1. State : optimal p를 current state의 CSI g\_t로만 찾기는 힘들어서 c, p라는 보충 요소 등장



1. Action : DQN의 action은 discrete해야하기 때문에 A-1 level로 나누어줌



1. Reward : 정교하게 디자인해도 대부분이 suboptimal로 수렴하기 때문에 그냥 downlink를 바로 reward.



1. Deep neural network

* 4 layer feed-forward
* The number of neurons of 2 hidden layers = 128, 64
* Activation function of output = linear, of 2 hidden = ReLU

n : BS, k : user

Independent channel gain(CSI information)

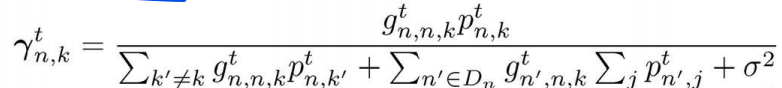
Small scale complex fading element

Large scale fading component

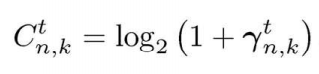
first-order complex Gauss-Markov process

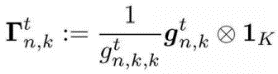


, where J0 = first kind zero-order Bessel function, f\_d = maximum doppler frequency, T\_s = time interval



D\_n = set of interference cells around the n-th cell. P = emitting power of BS, sigma = noise power

downlink rate of this link.



agent는 perfect CSI information가지고 있다고 가정 후, logarithmic normalized interferer set 정의, 1k is a vector filled with K ones.

Multiple Channel Access using Deep Reinforcement Learning for congested vehicular Networks

Cited by 2 times, [2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)](https://ieeexplore.ieee.org/xpl/conhome/9121635/proceeding)

1. Environment :

* self-experience-based CW adaptation algorithm employing DRL
* vehicle(agents) broadcast the safety packet using V2V communication and receive transmission results from a VANET. Consequently, vehicles learn to adjust the optimum CW
* DSRC(Dedicated Short Range Communication)의 multi-channel operation : CCH(control-channel), SCH(service-channel)으로 나누어지는데 각 interval = 50ms. 그래서 모든 vehicles은 100ms마다 safety packet 전송(in CCH interval ; CCHI), SCH interval ; SCHI에는 가장 가까운(지정된) vehicle(node)에 unicast ACK 전송. SCHI 때 target vehicle 선택 & transmit ACK

1. state : <CW, F, S>

CW : contention window

F : frequency value

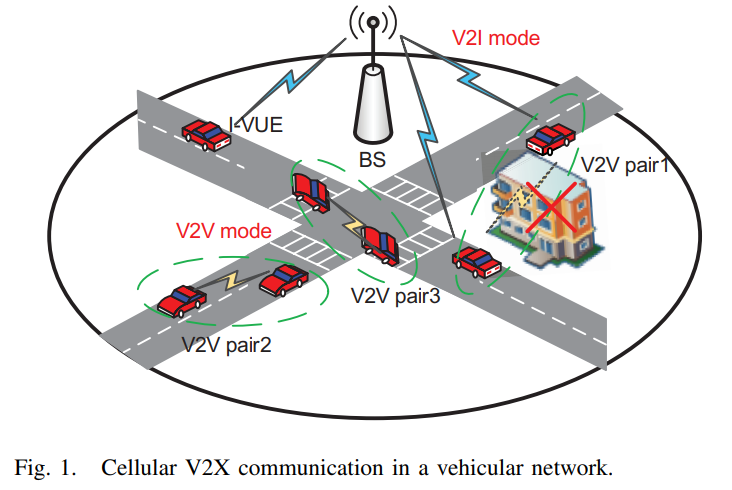
S : success rate

1. action : 3가지(keep, increase, decrease)
2. reward : broadcast 성공시 +1, 실패시 -1 -> 그래서 SCHI 때 각 agent의 broadcast성공인지 판단 가능
3. Deep neural network

* 3 hidden layers with the number of neurons 256, 128, 64 using Leaky-Relu as activation function.

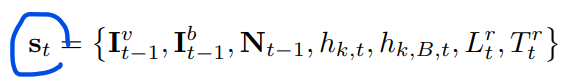
DQELR : An adaptive Deep Q-Network-based energy- and latency-aware routing protocol design for underwater acoustic sensor networks.

Cited by 33 times, [IEEE Access](https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6287639) ( Volume: 7)

1. Environment

BS : 중앙 위치, VUE : random distributed, single antenna

1. State



I\_v : the received interference power at the V2V receiver

I\_b : the received interference power at the BS

N : the number of selected neighbors

h\_k,t : the large scale channel gain from the V2V transmitter to its corresponding V2V receiver

h\_k,B,t : the large scale channel gain from the V2V transmitter to the BS

L : current load

T : remaining time to meet the latency threshold

1. Action

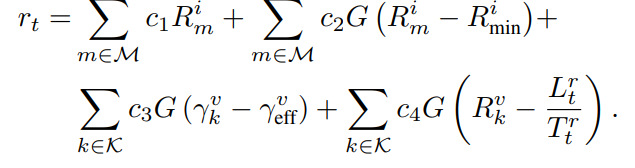


A : the RB allocation

S : communication mode selection

P : transmit power level of the V2V transmitter

1. Reward



1st : sum capacity revenue of I-VUEs

2nd : penalty of unsatisfied capacity for I-VUEs

3rd, 4th : impacts of the reliability and latency requirement

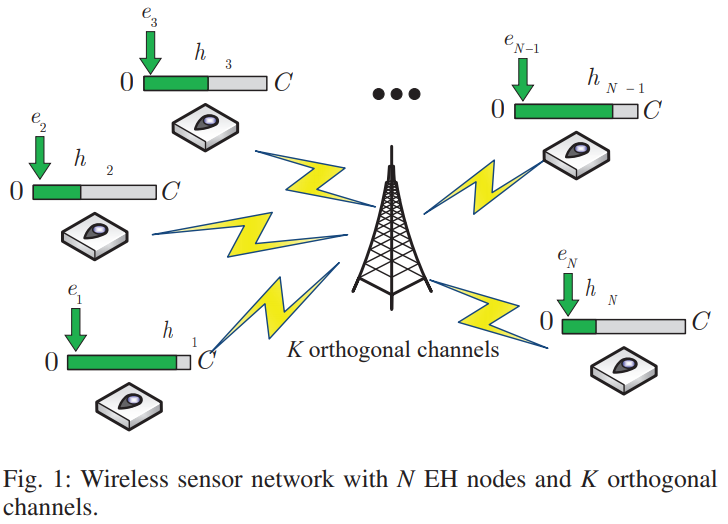
1. Deep neural network

* 1 hidden layer(256), 모두 fully connected layer,

Partially Observable Double DQN Based IoT Scheduling for Energy Harvesting

Cited by 3 times [2019 IEEE International Conference on Communications Workshops (ICC Workshops)](https://ieeexplore.ieee.org/xpl/conhome/8751668/proceeding)

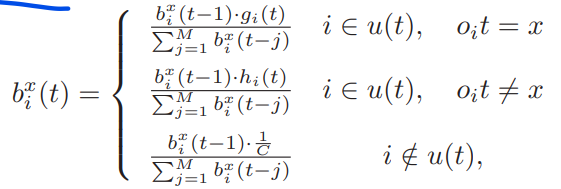
1. Environment

POMDP라 가운데 BS가 node의 상태를 완전히 알 수 없다.

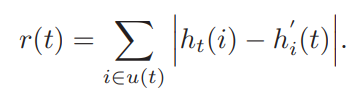
1. State

MDP라면 전체 정보를 알기에 이렇게 정의를 하겠지만,,



POMDP이기에 b = belief state 정의

1. Action : scheduling policy(broadcast to all nodes and receive the information about current power)
2. Reward

h : the power information about the scheduled node

h` : the residual power to BS again after attempting to transit data

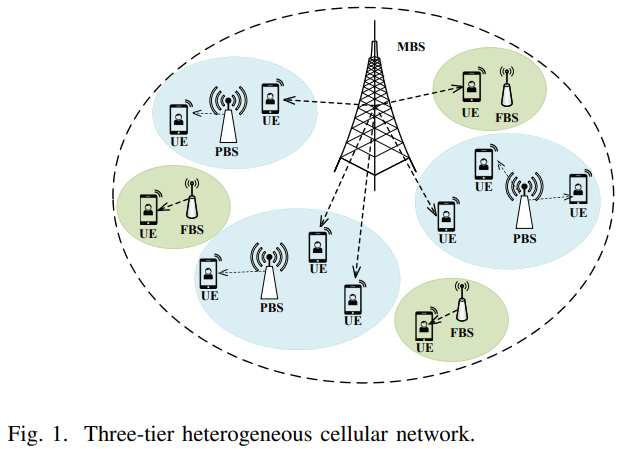
1. Deep neural network(keras in tensorflow)

3 fully connected layer(Sequential), hidden layer의 뉴론 수는 (input 뉴론 + out 뉴론) / 2, ReLU

Deep Reinforcement Learning for User Association and Resource Allocation in Heterogeneous Cellular Networks

Cited by 74 times [IEEE Transactions on Wireless Communications](https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=7693) ( Volume: 18, [Issue: 11](https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=8894903), Nov. 2019)

1. Environment

m개 MBS, p개 PBS, f개 FBS(small but high quality), N 랜덤 배치 유저

1. State
2. Action
3. Reward
4. Deep neural network