TWO-DIMENTIONAL ANTI-JAMMING COMMUNICATION BASED ON DEEP REINFORCEMENT LEARNING

Guoan Han[1]; Liang Xiao[1]; H. Vincent Poor[2]

[1]Dept. Communication Engineering, Xiamen University, China, lxiao@xmu.edu.cn [2]Dept. Electrical Engineering, Princeton University, USA, poor@princeton.edu

Abstract

In this paper, a two-dimensional anti-jamming communication scheme for cognitive radio networks is developed, in which a secondary user (SU) exploits both spread spectrum and user mobility to address jamming attacks, while not interfering with primary users (PUs). By applying a deep Q-network algorithm, this scheme determines whether to recommend that the SU leave an area of heavy jamming and chooses a frequency hopping pattern to defeat smart jammers. Without knowing the jamming model and the radio channel model, the SU derives an optimal anti-jamming communication policy using Q-learning in a proposed dynamic game, and applies a deep convolution neural network to accelerate the learning speed with a large number of frequency channels. The proposed scheme can increase the signal-to-interference-plus-noise ratio and improve the utility of the SU against cooperative jamming, compared with a Q-learning-only based benchmark system.

System Model

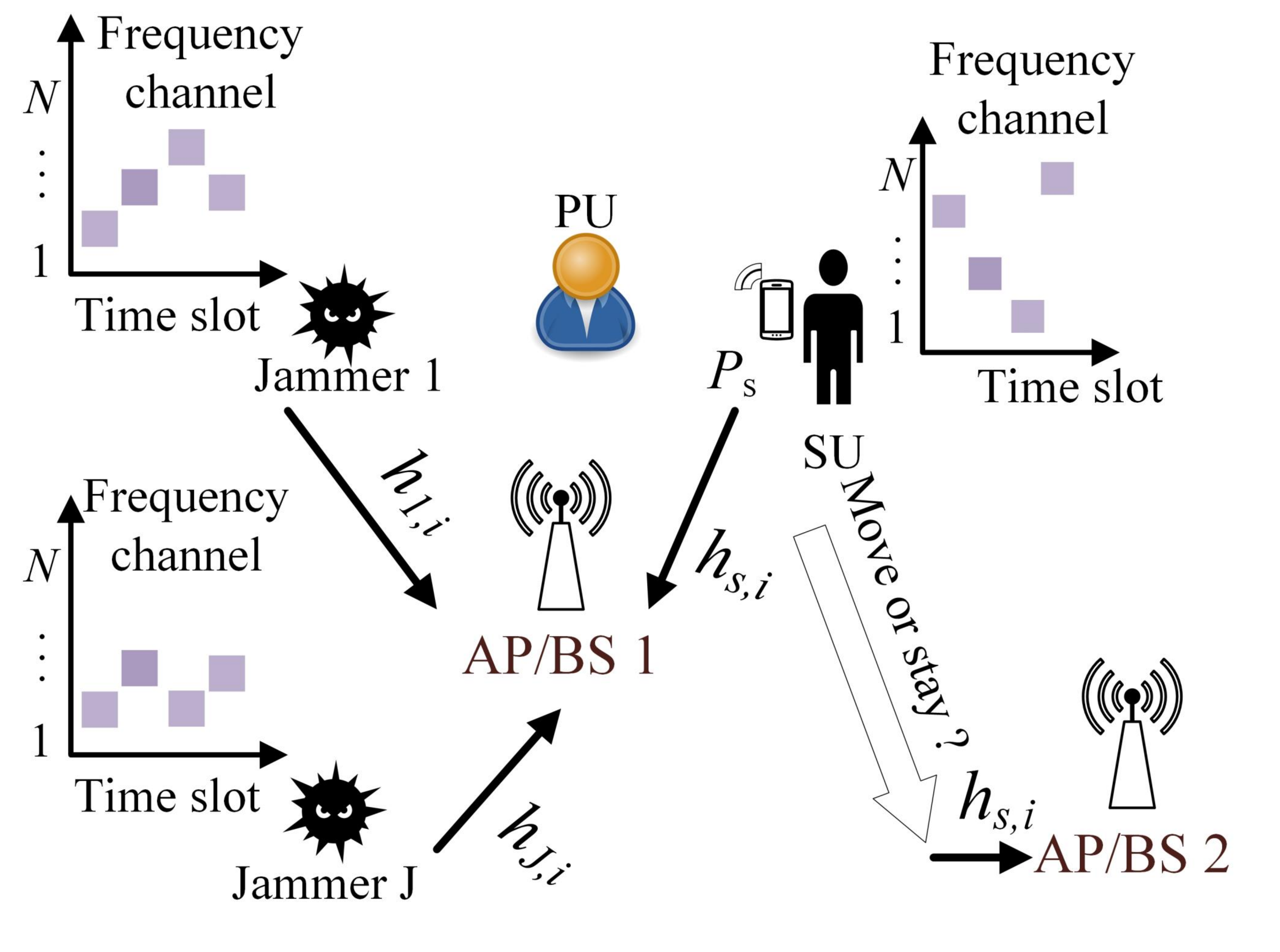

Figure 1. System model.

- SU chooses a communication channel and determines whether to leave the area
- J smart jammers block communication channels cooperatively

Game Model

The repeated interactions among an SU and jammers are formulated as a zero-sum dynamic anti-jamming communication game:

- The SU chooses an action $x_k \in \{0,1,\cdots,N\}$, if $x_k = 0$, the SU connects to a new AP/BS; otherwise, the SU uses channel x_k to send signals
- J cooperative jammers randomly located in the CRN and choose their jamming channels
- Both the SU and J jammers should have to avoid interfering with the PU

Utility function:

$$u_k(x,\mathbf{y}) = \frac{P_s h_{s,x} \lambda_k}{\sigma + \sum_{j=1}^{J} P_j h_{j,y_j} f(x = y_j)} - C_m f(x=0)$$

Presence of PUs; SINR of signals; Cost of defense

σ is the receiver noise power, and $f(\xi)$ is an indicator function that equals 1 if ξ is true, and 0 otherwise.

System state:

Consist of the presence of PUs and the SINR of the signals at last time, i.e., $\mathbf{s}_k = [\lambda_{k-1}, \mathrm{SINR}_{k-1}]$

Method

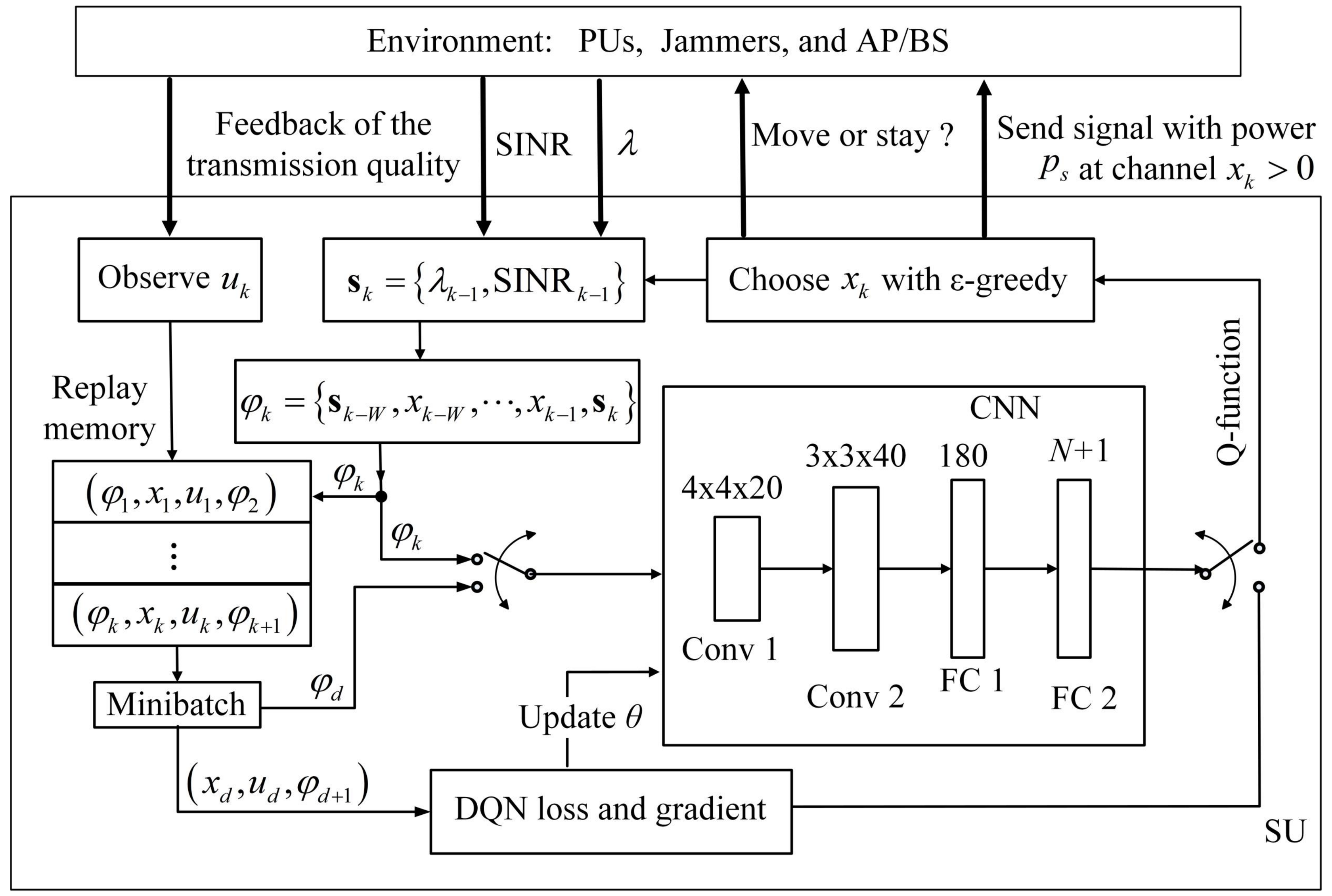

Figure 2. DQN-based 2-D anti-jamming communication system.

Challenges:

- Both the jamming model and the radio channel model are unware to the SU
- The size of available system state set is large
- The size of feasible frequency channels is large

Q-learning:

- Learning with incomplete information via trials
- Choose an action based on the Q-function, which is the expected discounted long-term utility for state $\mathbf{s}$ and action x
- Suffer from the curse of high-dimensional

$$Q(\mathbf{s}_k, x_k) = (1-\alpha)Q(\mathbf{s}_k, x_k) + \alpha(u_k + \gamma \max_{0 \le x' \le N} Q(\mathbf{s}_{k+1}, x'))$$

Deep Q-network :

- Combine reinforcement learning and deep learning
- Address the curse of high-dimensional of Q-learning by using convolutional neural network
- Significant improvements in the learning speed

$$L(\theta_k) = \mathrm{E}\left[\left(u + \gamma \max_{0 \le x' \le N} Q(\varphi', x'; \theta_{k-1}) - Q(\varphi, x; \theta_k)\right)^2\right]$$

Simulation results

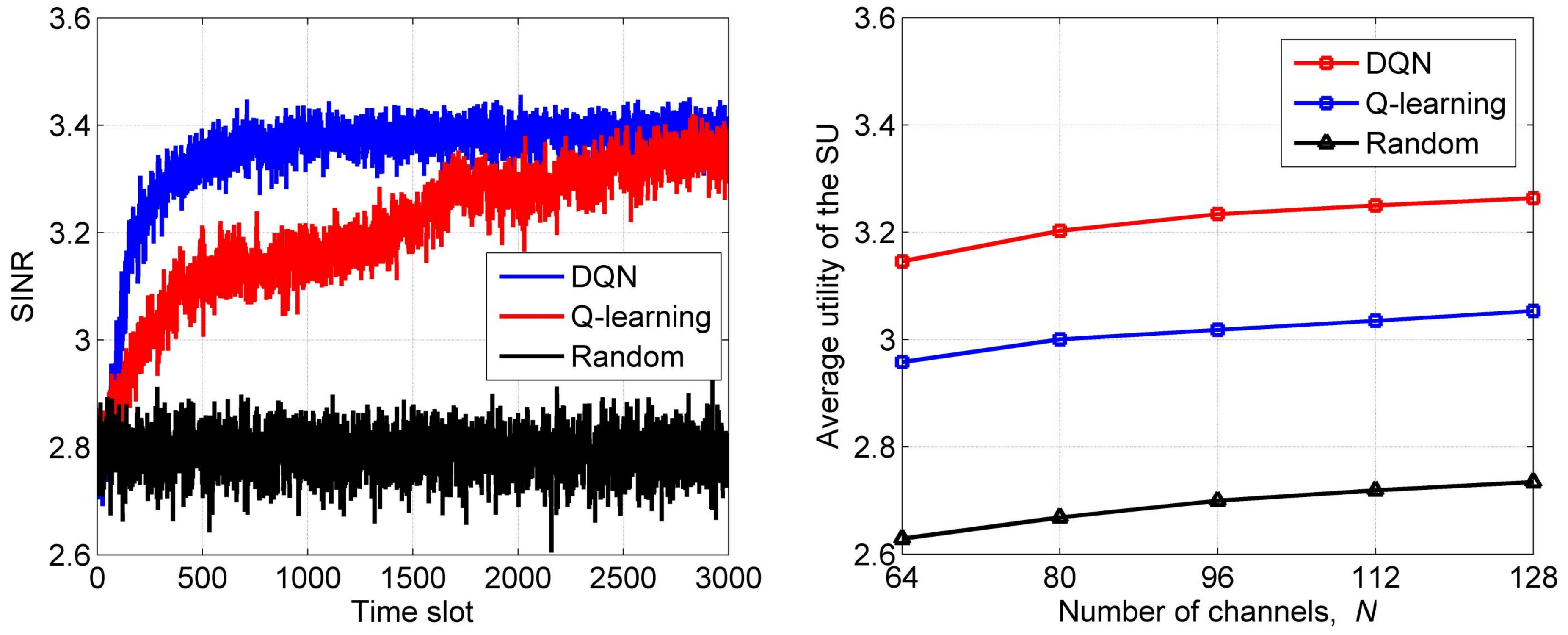

Figure 3. Performance of the DQN-based 2-D anti-jamming scheme.

Conclusions

- We have formulated a dynamic anti-jamming communication game for CRNs, which exploits both **frequency hopping and user mobility** to improve the SINR of the signals against cooperative smart jammers
- A DQN-based 2-D communication system is proposed for an SU to achieve the optimal anti-jamming policy
- The proposed 2-D DQN-based anti-jamming system outperforms the Q-learning strategy with **a faster convergence rate**, higher SINR, lower cost of defense and higher utility of the SU

References

1. L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2578-2590, Dec. 2015.
2. V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, and G. Ostrovski., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529-33, Jan. 2015.