

ADAPTIVE SPECTRUM MANAGEMENT IN COGNITIVE RADIO NETWORKS BASED ON Q-LEARNING AND MARKOV OPTIMIZATION

Merkulenko Y.S., Savchenko M.V.

National Technical University "Kharkiv Polytechnic Institute", Kharkiv

The paper investigates the integration of Markov models with machine learning algorithms for solving dynamic spectrum access problems in cognitive radio networks [1]. The main focus is on a decentralized approach, in which each node makes decisions on the use of spectral resources independently, based on local information and the principles of adaptive learning. The modeling is based on the Markov Decision Process (MDP), which reflects the functioning of the system under conditions of stochastic changes in spectrum availability [2].

The paper considers the mechanisms of combining Markov models with reinforcement learning methods, in particular the Q-learning algorithm. It is shown how cognitive network nodes are able to learn effective access strategies, receiving positive rewards for rational spectrum use and penalties for violating the rights of primary users. The characteristics of the changing environment, which is modeled through probabilistic transitions between the "busy" and "free" states, are also analyzed.

The proposed approach proved to be effective in the tasks of increasing throughput and reducing interference in conditions of limited spectral resources. The results obtained show that the use of learning algorithms built on the basis of Markov models allows achieving a balanced combination between the search for new spectral opportunities and the efficient use of existing resources. This creates the prerequisites for the development of more autonomous and flexible cognitive radio systems.

The combination of Q-learning with Markov modeling has shown high efficiency in solving problems of dynamic access to the spectrum in decentralized conditions. Compared to classical methods, the proposed approach provides more efficient use of the available spectrum and reduction of interference due to the ability of the system to adapt in conditions of uncertainty.

Additionally, the analysis indicates the importance of tuning the learning parameters, in particular the learning rate and the discount factor, which significantly affect the overall performance. Choosing the optimal values of these parameters allows for a balance between rapid adaptation to changes in the environment and stable system operation in the long term, which makes this approach promising for application in modern wireless networks.

References:

1. Pustovoitov, P., Okhrimenko, M., Voronets, V., & Udalov, D. (2021). The speed calculating increasing method of the markov model network node. *Advanced Information Systems*, 5(3), 13–17. <https://doi.org/10.20998/2522-9052.2021.3.02>.
2. Pustovoitov, P., Voronets, V., Voronets, O., Sokol, H., Okhrymenko, M. (2024). Assessment of QOS indicators of a network with UDP and TCP traffic under a node peak load mode. *Eastern-European Journal of Enterprise Technologies*, 1 (4 (127)), 23–31. DOI: <https://doi.org/10.15587/1729-4061.2024.299124>.