

*Пятін І.С., к.т.н., доцент,
Хмельницький політехнічний фаховий коледж
Бойко Ю.М., д.т.н., професор,
Хмельницький національний університет*

RL МОБІЛЬНІ ТЕЛЕКОМУНІКАЦІЇ З LDPC КОДУВАННЯМ

Сучасні бездротові системи зв'язку працюють в умовах завмирань, завад, швидких змін стану каналу та зростаючих вимог до швидкості передавання і якості обслуговування. Використання адаптивної модуляції і кодування дає змогу динамічно змінювати вид модуляції та кодову швидкість відповідно до поточного стану каналу, забезпечуючи компроміс між спектральною ефективністю, часткою помилкових блоків і затримкою. Reinforcement Learning (рис. 1) здійснюється через багаторазову взаємодію із середовищем і є інтелектуальною надбудовою для вибору режиму передавання. Замість жорсткої таблиці порогів для певного відношення сигнал/шум присутній агент, який після кожного фрейму або інтервалу передавання спостерігає стан каналу, вибирає режим передавання і отримує “винагороду”.



Рис. 1. а – Схема взаємодії агента RL з середовищем; б – схема системи зв'язку, що використовує навчання з підкріпленням (RL)

Агентом є блок керування адаптацією на боці передавача. Саме він вирішує, який режим роботи вибрати для наступного фрейму. Середовище включає канал, демодулятор, LDPC-декодер, інформація про якість якого передається через CQI, HARQ. Стан – це вектор ознак: оцінене SNR; CQI; попередній режим роботи; останній ACK/NACK; кількість ітерацій декодування (затримка). Дія – це вибір сигнально-кодової конструкції (тип модуляції, кодова швидкість).

Винагорода повинна змушувати агента знаходити компроміс між пропускну здатністю, BER і затримкою.

Для систем з Reinforcement Learning (RL) для адаптації лінії зв'язку критичні затримка обчислення і застарівання CQI, бо якщо агент працює повільно, канал швидко змінюється. Для FPGA-платформи AXU2CGB, тракт передавання та приймання реалізується у програмованій логіці, а RL-агент – на процесорній системі.

Адаптивну систему зв'язку з RL представляємо як задачу “стан → вибір режиму роботи → винагорода → оновлення правил”. Передача індексу режиму роботи до приймача відбувається каналом управління.

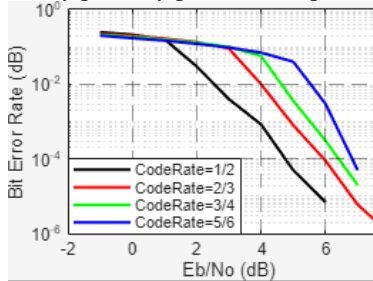


Рис. 2. – Залежність BER від Eb/No для різних CodeRate

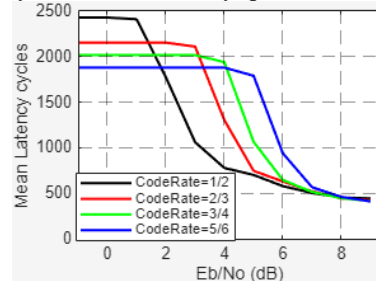


Рис. 3. – Залежність середньої затримки від Eb/No для різних CodeRate

LDPC-декодер працює у режимі з early termination: при малому Eb/No декодер працює з максимальною кількістю ітерацій, а при зростанні Eb/No зупиняється раніше, тому середня затримка падає. Виходячи з отриманих залежностей, одна ітерація триває 284 системні такти, а базова складова алгоритму декодування триває 157 тактів.

За діапазонами Eb/No (рис. 2-3) можна зробити такий висновок: при малих Eb/No декодер працює в режимі насичення і використовує максимальні 8 ітерацій. У діапазоні 2–4 dB маємо перехідний режим: середня затримка швидко зменшується, а розкид між фреймами найбільший. Для 6–7 dB декодер завершує декодування за 1–2 ітерації, затримка зменшується більш ніж у 5 разів.

Використання Reinforcement Learning корисно в області Eb/No \approx 3...6 dB, де спостерігається розкид затримки та кількості ітерацій між фреймами, що дозволяє використовувати орієнтоване на фреймах адаптивне керування параметрами декодування. При низькому BER Reinforcement Learning можна використовувати для адаптивного зменшення обчислювальної складності декодування.

Список використаних джерел:

1. Xie, C., El-Hajjar, M., Ng, S.X.: Machine learning assisted adaptive LDPC coded system design and analysis. IET Commun.18, 1–10 (2024). <https://doi.org/10.1049/cmu2.12707>