

КООПЕРАТИВНЕ НАВЧАННЯ З ПІДКРІПЛЕННЯМ В ІНТЕЛЕКТУАЛЬНИХ СИСТЕМАХ

Мацішевський В. О.

Науковий керівник – доц., к. т. н. Месюра В. І.

Під інтелектуальною системою розуміють таку систему, в якій для вирішення поставленої задачі використовуються методи імітації розумової діяльності людини. На сьогоднішній день, методи штучного інтелекту успішно застосовуються у багатьох сферах людської діяльності. Однією із перспективних і поки що мало досліджених областей є застосування інтелектуальних методів для розв'язання задачі маршрутизації пакетів у мобільних бездротових комп'ютерних мережах. У таких мережах кожен вузол виконує водночас дві ролі: виробника/споживача потоків пакетів даних і маршрутизатора пакетів, призначених для інших вузлів. При цьому вузли можуть вільно пересуватись у середовищі, властивості якого заздалегідь визначити неможливо.

На сьогоднішній день для розв'язання поставленої задачі використовується три групи протоколів маршрутизації: проактивні, реактивні і гібридні. Кожен існуючих протоколів має свої переваги і недоліки, однак жоден із них не відображає суть і головну відмінність мобільних комп'ютерних мереж від класичних мереж: мобільна мережа є системою, яку важко описати і яка постійно змінюється, а тому використання у ній строгих класичних (не інтелектуальних) алгоритмів є малоефективним. Однак для розв'язання задач у подібних системах можна успішно використовувати інтелектуальні методи, у чому і полягає суть даної роботи.

У роботі процес маршрутизації розглядається як процес навчання інтелектуальних агентів (пакетів). Сам процес навчання базується на основі отриманих агентами «підкріплень» (відповідно більш успішні дії агентів отримують більші підкріплення). Математично процес описується за допомогою рівнянь Белмана і так званих Q-значень, які обчислюються кожен раз, коли пакет (агент) передається з деякого вузла s на сусідній вузол s' :

$$Q(a, s) \leftarrow Q(a, s) + \alpha \cdot (R(s) + \gamma \cdot \max_{s'} Q(a', s') - Q(a, s)),$$

де $Q(a, s)$ – вартість виконання дії a в стані s ;

α – функція швидкості навчання;

$R(s)$ – функція винагороди;

γ – коефіцієнт знецінення.

Такий підхід до маршрутизації пакетів, оснований на навчанні з підкріпленням, дозволив побудувати стійку до постійних змін структури і властивостей мережі і водночас просту у реалізації систему маршрутизації.