

ВЫБОР ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ КАК ГИПЕРПАРАМЕТРА ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ ПРИ ПОИСКЕ ОПТИМАЛЬНОГО ПУТИ

Москалев П.В., Стебулянин М.М., Мягков А.С.

ФГАОУ ВО «МГТУ «СТАНКИН», РФ, 127055, Москва, Вадковский пер. 1,
moskaleff@mail.ru, mmsteb@rambler.ru, almyagkov99@yandex.ru

При поиске оптимального пути мобильного робота в физическом пространстве с препятствиями последнее может быть представлено в виде решетки. Тогда шаг h этой решетки будет гиперпараметром, определяющим область допустимых решений. Предложенный алгоритм можно рассматривать как разновидность метода обучения с подкреплением. На первом этапе с помощью перколяционного алгоритма Лиса выполняется анализ достижимости, что эквивалентно верификации связности пространства состояний при заданной дискретизации h . На втором этапе с помощью модифицированного алгоритма A^* реализуется поиск оптимального пути, где функция стоимости пути определяется эвристикой, прогнозирующей стоимость состояния, и штрафом за вход в зону безопасности, моделирующим негативное вознаграждение за рискованные участки пути. На третьем этапе с помощью комбинации алгоритмов Гилберта-Джонсона-Кирти и расширяющегося многогранника реализуется механизм уточнения ограничений, обеспечивающих безопасность решений, найденных на дискретной решетке, в непрерывном физическом пространстве.

Статистическое моделирование позволило выявить нелинейную зависимость вероятности существования допустимого решения $P(q|h)$ от гиперпараметра h и безразмерной плотности препятствий q . Успешная аппроксимация этой зависимости логистической функцией указывает на наличие критического значения $h_c(q)$, подобного порогу протекания в теории перколяции, за которым среда для агента становится «непроходимой». Эффективность найденного решения $E(q|h)$ демонстрирует насыщение при малых h , что соответствует известному в обучении с подкреплением принципу «убывающей отдачи» с ростом сложности модели: после определенного порога дальнейшая детализация мало улучшает качество решения, но сильно увеличивает сложность его поиска. Тогда критическое значение шага $h_c(q)$ можно рассматривать как точку Парето-оптима в пространстве «связность среды — сложность обучения».

Таким образом, шаг решетки является гиперпараметром, определяющим свойства процесса принятия решений при управлении мобильным роботом. Оптимизация этого параметра должна предшествовать процессу обучения. Найденные в [1] аппроксимации $P(q|h)$ и $E(q|h)$ создают базис для разработки алгоритмов обучения с подкреплением, адаптивно выбирающих оптимальное разрешение в задачах управления.

Работа выполнена при финансовой поддержке Минобрнауки РФ (госзадание в области научной деятельности FSFS-2024-0012).

Литература

1. Москалев П.В., Стебулянин М.М., Мягков А.С. Влияние пространственного разрешения на оптимальность пути мобильного робота в двумерных решеточных моделях // *Компьютерные исследования и моделирование*. Т. 17, № 6. 2025 [в печати].