

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 004.852 + 004.896

DOI: 10.17212/2782-2001-2023-4-55-68

Нейросетевой метод планирования пути в двумерном пространстве^{*}

Д.С. ЛУКИН^а, Е.Ю. КОСЕНКО^б

347928, Ростовская область, г. Таганрог, пер. Некрасовский, 44, Институт радиотехнических систем и управления ЮФУ

^аdlukin@sfedu.ru, ^бekosenko@sfedu.ru

В настоящее время роботизация различных сфер человеческой жизни движется высокими темпами. Роботы различного типа и назначения применяются повсеместно, от роботов-кладовщиков и роботов-инвентаризаторов, двигающихся по заданному маршруту или маркерам, до высокотехнологичных робототехнических комплексов, решающих сложные задачи с минимальным участием оператора. Робототехника продолжает развиваться, и ее потенциал для автоматизации и решения различных задач постоянно расширяется. Одним из ключевых вопросов повышения автономности мобильных роботов является разработка новых и совершенствование существующих подходов к управлению движением роботов, в частности к планированию пути. В настоящей работе задача планирования пути решается с применением искусственных нейронных сетей и глубокого машинного обучения с подкреплением, при котором робот учится выбирать действия в среде таким образом, чтобы максимизировать некоторую численную награду или достичь определенной цели. Такой подход позволяет обеспечить планирование траектории движения путем моделирования среды, поведения робота, а также взаимодействия между ними. Метод обучения с подкреплением предоставляет эффективный способ для роботов и автономных систем обучаться адаптироваться к разнообразным условиям и выполнять задачи планирования пути. В настоящей работе исследуется возможность решения задачи планирования движения в заданную точку с применением метода приближенной оптимизации стратегии и метода «действие – критика». Полученные результаты показывают возможность решения поставленной задачи при обучении на относительно небольшом количестве эпизодов. Предложенный подход может применяться для управления наземными роботизированными системами различного назначения.

Ключевые слова: планирование пути, нейросетевой метод, обучение с подкреплением, движение робота, машинное обучение, метод PPO, метод A*, стратегия действий

^{*} Статья получена 07 июля 2023 г.

ВВЕДЕНИЕ

Планирование траектории мобильного робота – это процесс определения оптимального по заданному критерию пути от начальной точки до конечной точки в заданной среде или частично неопределенной среде. Этот процесс включает в себя анализ окружающей среды, построение карты окружающей среды, выбор и применение алгоритма планирования пути движения робота. Планирование траектории является важным компонентом, требующим достижения высокой точности, эффективности и безопасности при управлении движением мобильных роботов в различных средах. Существует множество алгоритмов для выполнения этой задачи [1], каждый из которых имеет свои преимущества и недостатки в различных ситуациях. Рассмотрим некоторые из основных алгоритмов планирования пути.

Алгоритмы дискретного поиска, такие как A^* [2], алгоритм Дейкстры [2] и Θ^* [3], находят кратчайший путь между начальной и конечной точками, в том числе используя различные эвристики. Так, в работе [4] авторы используют алгоритм A^* для решения задачи планирования движения в условиях склада. Пространство представлено в виде сетки ячеек, переход из одной ячейки в другую имеет некоторую стоимость, увеличивающуюся с каждой итерацией алгоритма. На каждом шаге вычисляется оценочная функция

$$f(x, y, x_0, y_0) = g(x, y) + h(x_0, y_0),$$

где $g(x, y)$ – стоимость достижения начальной ячейки на данном шаге, $h(x_0, y_0)$ – эвристическая функция, оценивающая стоимость достижения заданной ячейки от начального положения и рассчитываемая по формуле

$$h(x_0, y_0) = |x_0 - x_{\text{target}}| + |y_0 - y_{\text{target}}|,$$

где x_0, y_0 – координаты начала пути; $x_{\text{target}}, y_{\text{target}}$ – координаты цели.

К общим недостаткам алгоритмов дискретного поиска кратчайшего пути можно отнести малую скорость обработки и повышенные требования к ресурсам для больших карт, а также малую эффективность для использования в динамических средах.

Метод потенциальных полей создает виртуальные потенциальные поля в окружающей среде, чтобы моделировать притяжение или отталкивание робота от различных объектов, таких как препятствия или целевые точки [5–7]. Данный метод является весьма популярным в робототехнике из-за простоты реализации, но в чистом виде имеет ряд недостатков, таких как проблема локальных минимумов, проблема узких проходов, неэффективность в стохастических средах. В связи с этими ограничениями использование метода потенциальных полей требует его модернизации или дополнительных алгоритмов. Например, в работе [8] авторы предлагают использовать метод виртуальной силовой гистограммы, основанный на уравнении Лапласа с использованием итерации Гаусса – Зейделя, для решения проблемы узких проходов и густых полос препятствий, а также метод гармонической потенциальной функции для устранения проблемы потенциальных ловушек.

Алгоритм *RRT* использует случайные деревья для поиска оптимального пути в сложных и непредсказуемых окружениях [9, 10]. Этот алгоритм также не лишен недостатков: медленная сходимость, неполная гарантированность нахождения оптимального решения, зависимость от выбора параметров. Часто алгоритм *RRT* комбинируют с другими алгоритмами. Так, в работе [11] автор предлагает совместное использование алгоритма *RRT* и метода потенциальных полей. Итоговая комбинация улучшает сходимость, но при этом увеличивает сложность вычислений.

Генетические алгоритмы используют эволюционный подход для поиска оптимального пути между начальной и конечной точками. Эволюционный подход – один из видов искусственного интеллекта, основная идея которого – имитировать процесс эволюции в природе: конкуренция, наследственность, мутации и отбор. За счет этого идет постепенное улучшение решений из поколения в поколение. Например, в работе [12] авторы использовали эволюционный подход для определения оптимального количества ног шагающего робота. С помощью предложенного метода авторам удалось найти компромисс между проходимостью робота и его габаритами. Генетические алгоритмы просто реализуемы и способны обходить локальные минимумы, но при этом не гарантируют получения оптимального решения и имеют медленную сходимость.

Выбор алгоритма планирования пути должен основываться на конкретной задаче и доступных ресурсах.

Одним из перспективных на данный момент методов планирования пути является нейросетевой метод, который заключается в использовании средств машинного обучения для построения модели планирования пути. Этот метод основан на обучении нейронной сети на основе примеров из опыта реального мира или симуляций, что позволяет создать модель, которая может предсказывать оптимальный путь для робота в заданных условиях.

Одним из преимуществ нейросетевого метода планирования пути является то, что он может учитывать множество параметров, таких как геометрия окружающей среды, скорость и ускорение робота, обход препятствий и т. д., что делает его более гибким и адаптивным, чем традиционные методы планирования пути.

Примером применения нейросетевого метода планирования пути является использование глубоких нейронных сетей для планирования пути автономных автомобилей [12–14]. В этом случае нейросеть обучается на большом количестве данных, включая карты дорог, информацию о трафике, погодные условия и другие факторы, которые могут влиять на планирование пути. После обучения нейросеть может генерировать оптимальный путь для автомобиля в режиме реального времени, учитывая текущие условия на дороге.

Однако нейросетевой метод планирования пути также имеет свои ограничения и недостатки, такие как необходимость большого количества данных для обучения, неопределенность и сложность интерпретации полученных результатов.

Нейросетевой метод основан на обучении с подкреплением. В настоящее время одним из активно развивающихся способов машинного обучения является способ обучения с подкреплением. Обучение с подкреплением – это раздел машинного обучения, в котором робот учится принимать решения в среде,

чтобы максимизировать некоторую числовую награду. В обучении с подкреплением робот наблюдает текущее состояние среды, выбирает действие, получает награду и переходит в новое состояние (рис. 1). Постепенно робот учится выбирать действия, которые приводят к наибольшей награде.

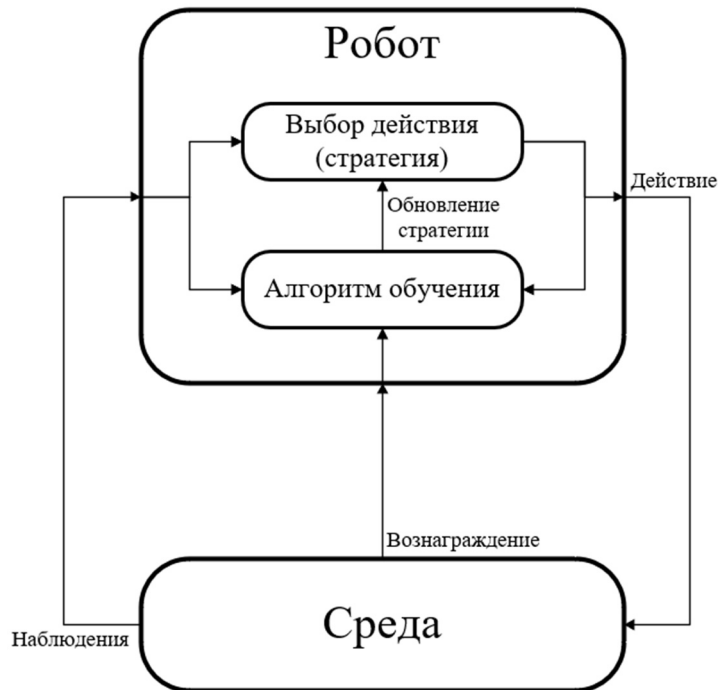


Рис. 1. Общая схема процесса обучения с подкреплением

Fig. 1. A general scheme of the learning process with reinforcement

Обучение с подкреплением широко используется для решения задач, в которых не существует явной формулировки правил и где объект должен самостоятельно научиться выбирать действия для достижения целей, используя метод проб и ошибок. Примерами таких задач могут быть игры [16], управление роботами [17–20], оптимизация рекламных кампаний и т. д.

В обучении с подкреплением используются различные алгоритмы, такие как Q -обучение [21, 22], SARSA [22, 23], TD -обучение [24], а также более сложные методы, которые используют нейронные сети для обработки информации о состоянии среды и выбора действий.

1. ПОСТАНОВКА ЗАДАЧИ

Робот,двигающийся в двумерном пространстве, должен достичь заданной неподвижной цели. Робот движется с постоянной скоростью $v = 1$ м/с в направлении угла φ относительно горизонтали. В течение одного шага робот может выбрать одно из трех возможных действий – изменить угол направления движения φ на -2° , 0° или 2° и переместиться в этом направлении.

Наблюдаемыми переменными являются координаты робота x, y , угол направления движения робота φ и угол между вектором скорости и направлением на цель α (рис. 2).

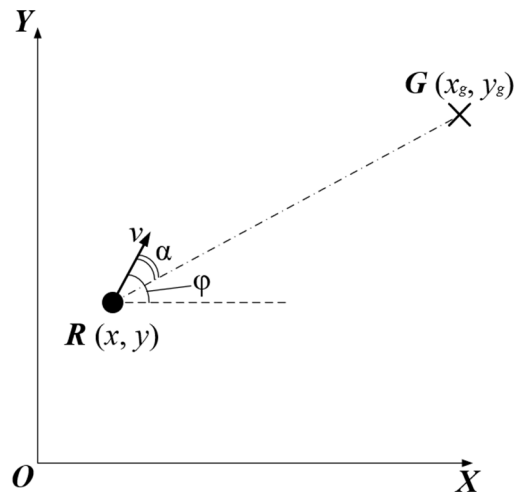


Рис. 2. Система координат среды и наблюдаемые углы

Fig. 2. The coordinate system of the medium and the observed angles

2. НАСТРОЙКА СРЕДЫ

В процессе обучения нейронной сети робот взаимодействует с окружающей средой, предпринимая действия, получая наблюдения и обновляя стратегию, – этот цикл называется эпизодом обучения. Эпизод заканчивается по заданному условию, в данном случае если робот вышел за пределы карты или достиг заданной цели.

В начале каждого эпизода (рис. 3) робот находится в определенном состоянии среды и должен принять решение о следующем действии на основе своей стратегии выбора действий. После выполнения действия робот получает информацию о состоянии среды (наблюдение) и подкрепление (награду – положительную или отрицательную). Этот процесс повторяется до достижения терминального состояния или определенного количества шагов.

Завершение каждого эпизода включает в себя обновление стратегии на основе полученного опыта, чтобы робот мог в будущем выбирать более оптимальные действия.

Положительную награду робот получает за правильно выдержанный угол направления ($|\alpha| \leq 2^\circ$) и отрицательную – за несоблюдение этого условия.

На рис. 3 изображена плоскость с роботом и целью в начале эпизода. Робот обозначен кружком в левой части плоскости. Для отслеживания направления движения робота на его пиктограмму добавлен отрезок, отображающий направление вектора скорости.

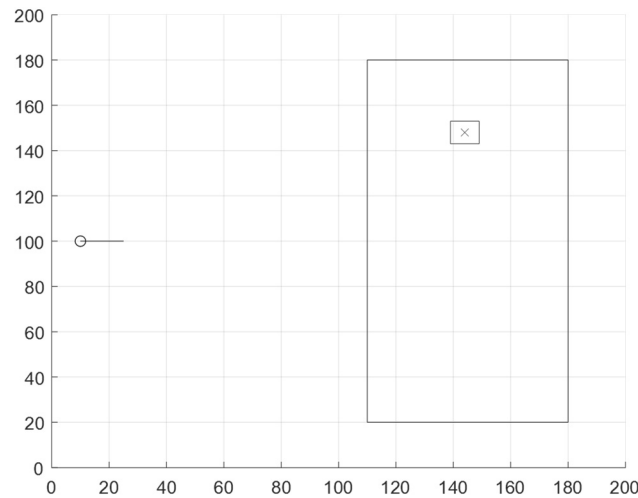


Рис. 3. Начало эпизода

Fig. 3. The beginning of the episode

Цель обозначена крестиком. Чтобы решить поставленную задачу планирования пути, робот должен достичь область вокруг цели, обозначенную прямоугольником.

Область возможного появления цели обозначена на плоскости большим прямоугольником.

В качестве метода обучения робота был выбран метод *PPO* (Proximal Policy Optimization) [25], который используется для обучения в задачах, связанных с принятием решений в динамических средах.

Метод *PPO* является алгоритмом, который основан на методе стохастической градиентной оптимизации и придерживается ограничения на изменение стратегии. Это позволяет избежать слишком больших изменений в стратегии, которые могут привести к нестабильности обучения.

В процессе обучения *PPO* использует опыт, полученный роботом во время взаимодействия со средой, для обновления параметров стратегии. Основная идея заключается в том, чтобы использовать опыт для оценки ожидаемой награды и обновлять параметры стратегии таким образом, чтобы максимизировать эту награду.

Оценку ожидаемой награды обеспечивает нейронная сеть, называемая критиком. Топология нейронной сети-критика данного робота изображена на рис. 5, а, она состоит из полносвязных слоев и слоев активации *ELU* (Exponential Linear Unit). Функция активации *ELU* описывается следующим уравнением [26]:

$$ELU(x) = \begin{cases} x, & \text{if } x > 0, \\ \alpha(e^x - 1), & \text{if } x \leq 0, \end{cases}$$

где α – параметр, регулирующий влияние отрицательных значений x . График функции активации *ELU* при $\alpha = 1$ изображен на рис. 4.

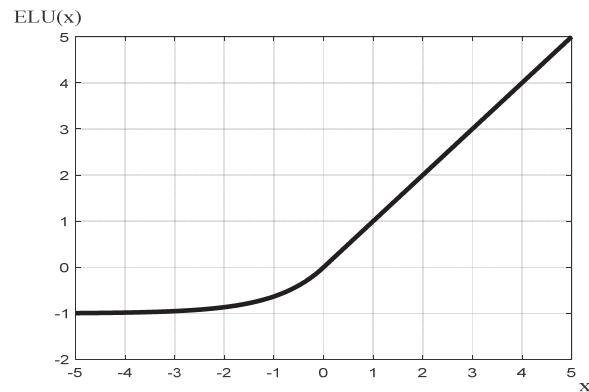


Рис. 4. График функции активации ELU при $\alpha = 1$

Fig. 4. A graph of the ELU activation function at $\alpha = 1$

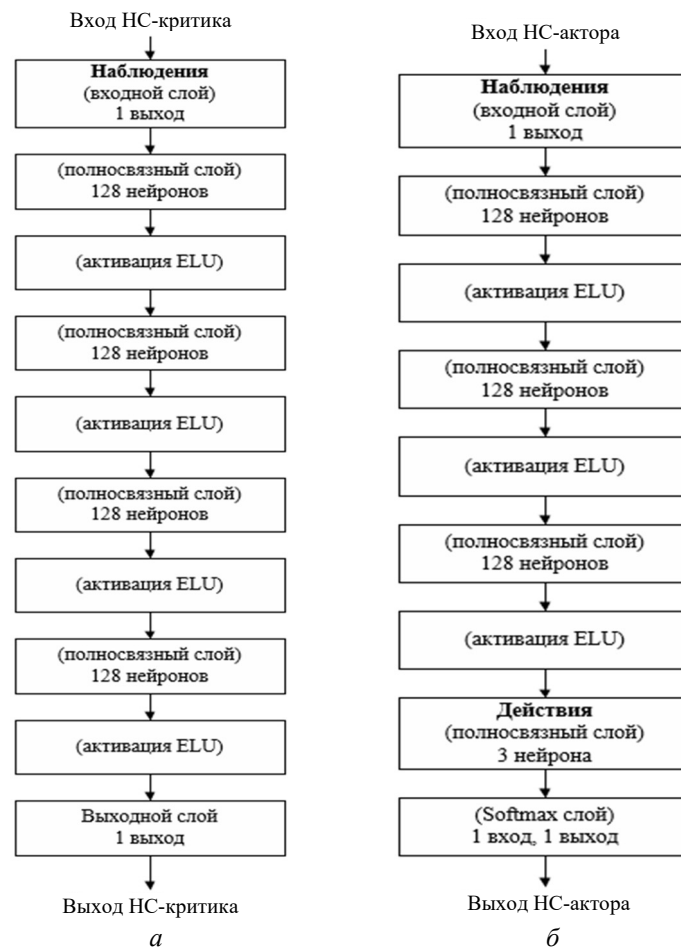


Рис. 5. Топология нейронной сети:

a – критика; b – актора

Fig. 5. Neural network topology:

a – is a critic; b – is an actor

Использование функции активации *ELU* способствует более быстрому обучению сети, чем функция активации *ReLU* (Rectified Linear Unit) [26].

Сами же действия предпринимаются другой нейронной сетью, называемой актором. Топология нейронной сети-актора представлена на рис. 5, б. Актор также состоит из полносвязных слоев и слоев активации *ELU*, но выходным слоем является слой активации *softmax*, функция которого выводит в качестве результата вектор, представляющий распределение вероятности списка потенциальных результатов.

В совокупности эти нейронные сети отвечают за поведение робота.

3. ОБУЧЕНИЕ И МОДЕЛИРОВАНИЕ

Робот обучался в течение 7000 эпизодов. На рис. 6 представлено несколько завершенных эпизодов моделирования обученной сети. Робот достигает случайно сгенерированной цели в заданном сегменте плоскости.

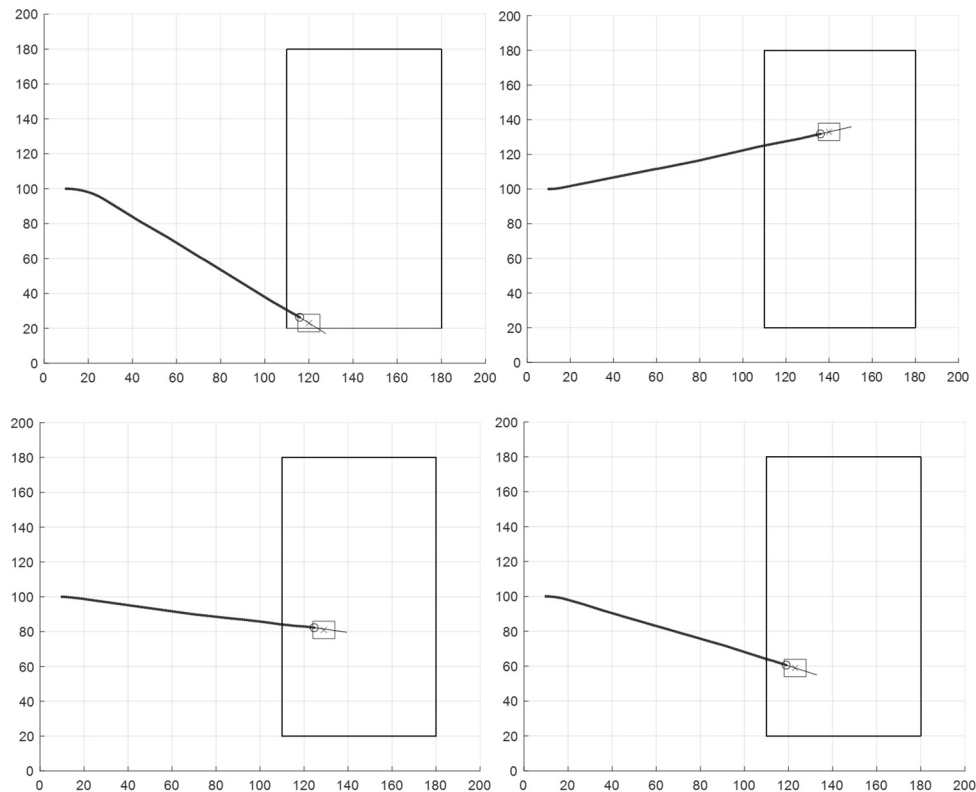


Рис. 6. Результаты моделирования обученного робота

Fig. 6. Simulation results of a trained robot

Сравним способность планирования пути обученной нейронной сети робота с алгоритмом A^* , реализованным в руководстве [27], по критериям оптимальности пути и времени выполнения. При одинаковых начальных позициях робота и цели алгоритм A^* выполняет задачу планирования пути за 37 с, в то время как нейронной сети робота понадобилось всего 2,6 с.

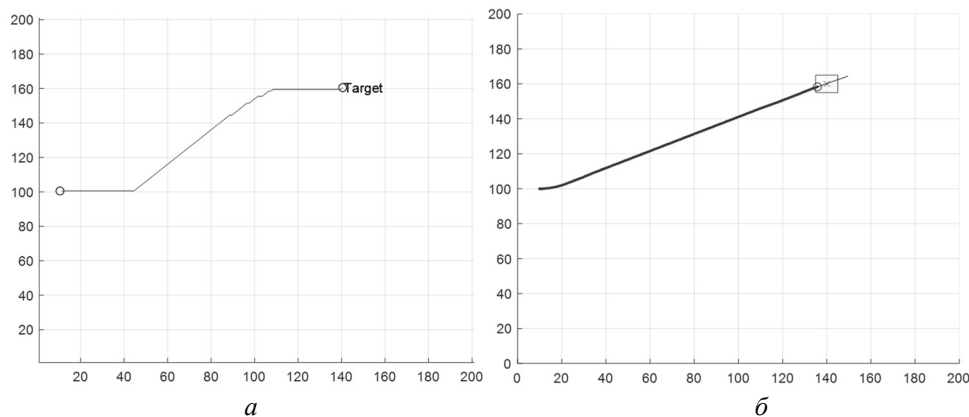


Рис. 7. Пройденный роботом путь при использовании:

a – алгоритма A*; b – нейросетевого метода

Fig. 7. The path traversed by the robot when using:

a – algorithm A*; b – neural network method

Нетрудно заметить, что пройденный роботом путь при использовании алгоритма A* (рис. 7, a) менее оптимален, чем при использовании нейросетевого метода (рис. 7, b). Это обусловлено особенностью работы алгоритма A*, движение робота может осуществляться только по ближайшим (к роботу) ячейкам сетки, т. е. минимальный шаг угла поворота робота составляет $|\varphi| = 45^\circ$, в то время как у обученного робота это значение составляет $|\varphi| = 2^\circ$ либо 0° за одно действие. Стоит отметить, что данное ограничение угла поворота робота можно избежать, используя непрерывное пространство состояний, однако это требует более детального описания среды.

ЗАКЛЮЧЕНИЕ

Таким образом, в настоящей работе представлен нейросетевой метод планирования пути, основанный на способе обучения с подкреплением и позволяющий научить модель робота в режиме реального времени выбирать кратчайший путь для достижения случайно заданной цели на основе получаемой обратной связи в виде вознаграждения или штрафа. Одним из ключевых преимуществ нейросетевого метода является способность обобщать поставленные задачи и применять их для решения новых ситуаций. После обучения на некотором наборе данных сеть может успешно решать схожие задачи без дополнительного обучения. Однако стоит отметить, что для более универсальной системы управления требуется большое количество данных. Несмотря на это, доступные на данный момент технологии и вычислительные ресурсы позволяют смоделировать практически любую ситуацию, необходимую для обучения.

В дальнейшем планируется усложнять среду путем введения большей стохастичности появления цели на карте, добавления препятствий, увеличения количества целей.

СПИСОК ЛИТЕРАТУРЫ

1. Казаков К.А., Семенов В.А. Обзор современных методов планирования движения // Труды Института системного программирования РАН. – 2016. – Т. 28 (4). – С. 241–294. – DOI: 10.15514/ISPRAS-2016-28(4)-14.
2. Алгоритмы поиска кратчайшего пути / А.Р. Худайбердина, Г.Ю. Беспялый, В.Е. Белай, А.В. Выграновская // Modern Science. – 2023. – № 1-2. – С. 144–148. – EDN EWHJYN.
3. Uras T., Koenig S. An empirical comparison of any-angle path-planning algorithms // Proceedings of the International Symposium on Combinatorial Search. – 2015. – Vol. 6 (1). – P. 206–210. – DOI: 10.1609/socs.v6i1.18382.
4. Бобырь М.В., Дородных А.А., Булатников В.А. Алгоритм поиска пути A* в задаче автоматизации процесса складирования изделий мехатронным комплексом // Мехатроника, автоматика и робототехника. – 2023. – № 11. – С. 151–154. – DOI: 10.26160/2541-8637-2023-11-151-154. – EDN PLILNX.
5. Платонов А.К., Карпов И.И., Кирильченко А.А. Метод потенциалов в задаче прокладки трассы // Препринты ИПМ им. М.В. Келдыша. – 1974. – № 124. – 27 с.
6. Khatib O. Real-time obstacle avoidance for manipulators and mobile robots // The International Journal of Robotics Research. – 1985. – Vol. 5 (1). – P. 500–505.
7. Филимонов А.Б., Филимонов Н.Б. Вопросы управления движением мобильных роботов методом потенциального наведения // Мехатроника, автоматизация, управление. – 2019. – Т. 20, № 11. – С. 677–685. – DOI: 10.17587/mau.20.677-685. – EDN KNYGTN.
8. Sabudin E.N., Omar R.B., Melor C.K. Potential field methods and their inherent approaches for path planning // ARPN Journal of Engineering and Applied Sciences. – 2016. – Vol. 11 (18). – P. 10801–10805.
9. Исследование методов планирования движения в двумерных картографированных средах / В.Х. Пшихопов, М.Ю. Медведев, Д.О. Бросалин, М.А. Васильева, Б.В. Гуренко, Н. Хамдан // Известия ЮФУ. Технические науки. – 2022. – № 3 (227). – С. 170–192. – DOI: 10.18522/2311-3103-2022-3-170-192. – EDN LDIEZZ.
10. Алгоритмы планирования траекторий в двумерной среде с препятствиями / В.Х. Пшихопов, М.Ю. Медведев, В.А. Костюков, Ф. Хуссейн, А. Кадим // Информатика и автоматизация. – 2022. – Т. 21, № 3. – С. 459–492. – DOI: 10.15622/ia.21.3.1. – EDN MVQDCS. – Яз. англ.
11. Чжан Ч. Исследование планирования пути улучшенного алгоритма RRT на основе искусственного потенциального поля // StudNet. – 2021. – Т. 4, № 6. – С. 1961–1977.
12. Буличев О.В., Малолетов А.В. Метод оптимизации количества ног шагающего робота на основе эволюционного алгоритма // Известия Волгоградского государственного технического университета. – 2022. – № 9 (268). – С. 12–19. – DOI: 10.35211/1990-5297-2022-9-268-12-19. – EDN DBEDNF.
13. Learning from experience for rapid generation of local car maneuvers / P. Kicki, T. Gawron, K. Ćwian, M. Ozay, P. Skrzypczyński // Engineering Applications of Artificial Intelligence. – 2021. – Vol. 105. – P. 104399. – DOI: 10.1016/j.engappai.2021.104399.
14. Path planning via an improved DQN-based learning policy / L. Lv, S. Zhang, D. Ding, Y. Wang // IEEE Access. – 2019. – Vol. 7. – P. 67319–67330. – DOI: 10.1109/ACCESS.2019.2918703.
15. Чугунов Р.А., Александрова Т.В. Применение нейронных сетей для навигации мобильного робота по видеоориентирам // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых, г. Томск, 7–11 ноября 2016 г.: в 2 т. – Томск: Изд-во ТПУ, 2016. – Т. 1. – С. 331–332.
16. Dota 2 with Large Scale Deep Reinforcement Learning / C. Berner [et al.]. – arXiv:1912.06680. – 2019.
17. Обучение с подкреплением в задачах управления беспилотными летательными аппаратами в среде умного города / Н.С. Власова, А.Н. Лебединец, Н.В. Быков, М.С. Товарнов // Будущее машиностроения России: Четырнадцатая Всероссийская конференция молодых ученых и специалистов: сборник докладов: в 2 т. – М., 2022. – Т. 2. – С. 103–110. – EDN PWUWDP.
18. Дудаков А.С., Турсунов Т.Р., Филимонов Н.Б. Метод глубокого обучения с подкреплением в задачах маршрутизации движения мобильных роботов в среде с препятствиями // Мехатроника, автоматика и робототехника. – 2023. – № 11. – С. 7–13. – DOI: 10.26160/2541-8637-2023-11-7-13. – EDN KDIJYO.
19. Parkhomenko V., Gayda T., Medvedev M. Intelligent system for countering groups of robots based on reinforcement learning technologies // Frontiers in robotics and electromechanics / ed. by

A. Ronzhin, V. Pshikhov. – Singapore: Springer, 2023. – P. 135–146. – (Smart Innovation, Systems and Technologies; vol. 329). – DOI: 10.1007/978-981-19-7685-8_9.

20. Нейросетевая система управления группой роботов в неопределенной двумерной среде / А.Р. Гайдук, О.В. Мартянов, М.Ю. Медведев, В.Х. Пшихопов, Н. Хамдан, А. Фархун // Мехатроника, автоматизация, управление. – 2020. – Т. 21, № 8. – С. 470–479. – DOI: 10.17587/mau.21.470-479. – EDN LAAABX. – Яз. англ.

21. Watkins C.J.C.H., Dayan P. Q-learning // Machine Learning. – 1992. – Vol. 8. – P. 279–292. – DOI: 10.1007/BF00992698.

22. Asynchronous methods for deep reinforcement learning / V. Mnih [et al.] // Proceedings of Machine Learning Research. – 2016. – Vol. 48. – P. 1928–1937.

23. Rummery G.A., Niranjan M. On-line Q-learning using connectionist systems. Technical Report CUED/F-INFENG/TR 166. – Cambridge University Engineering Department, 1994.

24. Sutton R.S., Barto A.G. Reinforcement learning: an introduction. – 2nd ed. – Cambridge: MIT Press, 2018. – 526 p.

25. Proximal policy optimization algorithms / J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov. – arXiv:1707.06347. – 2017.

26. Clevert D., Unterthiner T., Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs). – arXiv:1511.07289. – 2015.

27. Premakumar P. A* (A Star) search for path planning tutorial: MATLAB Central File Exchange. – 2023. – URL: <https://www.mathworks.com/matlabcentral/fileexchange/26248-a-a-star-search-for-path-planning-tutorial> (accessed: 28.11.2023).

Лукин Дмитрий Сергеевич, аспирант кафедры систем автоматического управления Института радиотехнических систем и управления ЮФУ. Основное направление научных исследований – искусственный интеллект, нейронные сети. Имеет 4 печатные работы. E-mail: dlukin@sfedu.ru

Косенко Евгений Юрьевич, кандидат технических наук, доцент кафедры систем автоматического управления Института радиотехнических систем и управления ЮФУ. Основное направление научных исследований – представление и обработка данных: методы обработки статистических данных, нечеткая логика. Имеет свыше 280 научных и учебно-методических работ. E-mail: ekosenko@sfedu.ru

Lukin Dmitry S., postgraduate student at the Department of Automatic Control Systems of the Institute of Radio Engineering Systems and Management in the Southern Federal University. The main direction of his scientific research is artificial intelligence and neural networks. He has more than 4 printed works. E-mail: dlukin@sfedu.ru

Kosenko Evgeny Yu., PhD (Eng.), associate professor at the Department of Automatic Control Systems of the Institute of Radio Engineering Systems and Management in the Southern Federal University. The main direction of his scientific research is data representation and processing: statistical data processing methods, fuzzy logic. He has published more than 280 scientific and educational works. E-mail: ekosenko@sfedu.ru

DOI: 10.17212/2782-2001-2023-4-55-68

A neural network method for path planning in a two-dimensional space*D.S. LUKIN^a, E.Yu. KOSENKO^b*Institute of Radio Engineering Systems and Management of SFU, 44 Nekrasovsk Lane, Taganrog, 347928, Russian Federation*^a *dluhin@sfnedu.ru* ^b *eykosenko@gmail.com***Abstract**

Currently, the robotization of various spheres of human life is moving at a high pace. Robots of various types and purposes are used everywhere, from storage robots moving along a given route or markers to high-tech robotic complexes that solve tasks with minimal operator participation. Robotics technology continues to evolve, and its potential for automation and solving various tasks is constantly expanding. One of the key issues of increasing the autonomy of mobile robots is the development of new and improvement of the existing approaches to controlling the movement of robots, in particular to path planning. In this paper, the task of path planning is solved using artificial neural networks and deep machine learning with reinforcement, in which the robot learns to choose actions in the environment in such a way as to maximize some numerical reward or achieve a certain goal. This approach allows you to plan the trajectory of movement by modeling the environment, the behavior of the robot, as well as the interaction between them. The reinforcement learning method provides an effective way for robots and autonomous systems to learn to adapt to diverse conditions and perform path planning tasks. In this paper, the possibility of solving the problem of planning movement to a given point using the method of approximate strategy optimization and the "Action – Criticism" method is investigated. The results obtained show the possibility of solving the task when learning on a relatively small number of episodes. The proposed approach can be used to control ground-based robotic systems for various purposes.

Keywords: path planning, neural network method, reinforcement learning, robot movement, machine learning, PO method, A* method, action strategy

REFERENCES

1. Kazakov K.A., Semenov V.A. Obzor sovremennykh metodov planirovaniya dvizheniya [An overview of modern methods for motion planning]. *Trudy Instituta sistemnogo programmirovaniya RAN = Proceedings of the Institute for System Programming of the RAS*, 2016, vol. 28 (4), pp. 241–294. DOI: 10.15514/ISPRAS-2016-28(4)-14.
2. Khudaiberdina A.R., Asfar D.B., Garmatii N.I., Rotanov I.A. Algoritmy poiska kratchaishego puti [Algorithms for finding the shortest path]. *Modern Science*, 2023, no. 1-2, pp. 144–148. (In Russian).
3. Uras T., Koenig S. An empirical comparison of any-angle path-planning algorithms. *Proceedings of the International Symposium on Combinatorial Search*, 2015, vol. 6 (1), pp. 206–210. DOI: 10.1609/socs.v6i1.18382.
4. Bobyr M.V., Dorodnyh A.A., Bulatnikov V.A. Algoritm poiska puti A* v zadache avtomatizatsii protsessa skladirovaniya izdelii mekhatronnym kompleksom [The path search algorithm A* in the problem of automation storage products process of a mechatronic complex]. *Mekhatronika, avtomatika i robototekhnika*, 2023, no. 11, pp. 151–154. DOI: 10.26160/2541-8637-2023-11-151-154. (In Russian).
5. Platonov A.K., Karpov I.I., Kiril'chenko A.A. Metod potentsialov v zadache prokladki trassy [The method of potentials in the task of laying a route]. *Preprinty IPM im. M.V. Keldysha = Keldysh Institute preprints*, 1974, no. 124. 27 p.

* Received 07 July 2023.

6. Khatib O. Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research*, 1985, vol. 5 (1), pp. 500–505.
7. Filimonov A.B., Filimonov N.B. Voprosy upravleniya dvizheniem mobil'nykh robotov metodom potentsial'nogo navedeniya [Issues of motion control of mobile robots based on the potential guidance method]. *Mekhatronika, avtomatizatsiya, upravlenie* = *Mechatronics, Automation, Control*, 2019, vol. 20 (11), pp. 677–685. DOI: 10.17587/mau.20.677-685.
8. Sabudin E.N., Omar R.B., Melor C.K. Potential field methods and their inherent approaches for path planning. *ARNP Journal of Engineering and Applied Sciences*, 2016, vol. 11 (18), pp. 10801–10805.
9. Pshikhopov V.Kh., Medvedev M.Yu., Brosalin D.O., Vasileva M.A., Gurenko B.V., Hamdan N. Issledovanie metodov planirovaniya dvizheniya v dvumernykh kartografirovannykh sredakh [Study of path planning methods in two-dimensional mapped environments]. *Izvestiya YFU = Izvestiya SFedU. Engineering sciences*, 2022, no. 3 (227), pp. 170–192. DOI: 10.18522/2311-3103-2022-3-170-192.
10. Pshikhopov V.Kh., Medvedev M.Yu., Kostjukov V., Houssein F., Kadhim A. Trajectory planning algorithms in two-dimensional environment with obstacles. *Informatika i avtomatizatsiya* = *Informatics and Automation*, 2022, vol. 21 (3), pp. 459–492. DOI: 10.15622/ia.21.3.1.
11. Zhang Ch. Issledovanie planirovaniya puti uluchshennogo algoritma RRT na osnove iskusstvennogo potentsial'nogo polya [Research on path planning of improved RRT algorithm based on artificial potential field]. *StudNet*, 2021, vol. 4 (6), pp. 1961–1977. (In Russian).
12. Bulichev O.V., Maloletov A.V. Metod optimizatsii kolichestva nog shagayushchego robota na osnove evolyutsionnogo algoritma [Optimizing the number of legs of a walking robot based on evolutionary algorithm]. *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta* = *Izvestia of Volgograd State Technical University*, 2022, no. 9 (268), pp. 12–19. DOI: 10.35211/1990-5297-2022-9-268-12-19.
13. Kicki P., Gawron T., Ćwian K., Ozay M., Skrzypczyński P. Learning from experience for rapid generation of local car maneuvers. *Engineering Applications of Artificial Intelligence*, 2021, vol. 105, p. 104399. DOI: 10.1016/j.engappai.2021.104399.
14. Lv L., Zhang S., Ding D., Wang Y. Path planning via an improved DQN-based learning policy. *IEEE Access*, 2019, vol. 7, pp. 67319–67330. DOI: 10.1109/ACCESS.2019.2918703.
15. Chugunov R.A., Aleksandrova T.V. [Application of neural networks for mobile robot navigation on video landmarks]. *Molodezh' i sovremennye informatsionnye tekhnologii* [Youth and modern information technologies]. Proceedings of the XIV International scientific and practical conference of students, postgraduate students and young scientists, Tomsk, 2016, vol. 1, pp. 331–332. (In Russian).
16. Berner C., et al. *Dota 2 with large scale deep reinforcement learning*. arXiv:1912.06680. 2019.
17. Vlasova N.S., Lebedinets A.N., Bykov N.V., Tovarnov M.S. [Reinforcement learning in the tasks of controlling unmanned aerial vehicles in a smart city environment]. *Budushchee mashinostroeniya Rossii* [The Future of Russian engineering]. Proceedings of the Fourteenth All-Russian Conference of Young Scientists and Specialists. Moscow, 2022, vol. 2, pp. 103–110. (In Russian).
18. Dudakov A.S., Tursunov T.R., Filimonov N.B. Metod glubokogo obucheniya s podkrepleniem v zadachakh marshrutizatsii dvizheniya mobil'nykh robotov v srede s prepyatstviyami [The method of deep reinforcement learning in motion planning problem of mobile robots in an environment with obstacles]. *Mekhatronika, avtomatika i robototekhnika*, 2023, no. 11, pp. 7–13. DOI: 10.26160/2541-8637-2023-11-7-13.
19. Parkhomenko V., Gayda T., Medvedev M. Intelligent system for countering groups of robots based on reinforcement learning technologies. *Frontiers in robotics and electromechanics*. Ed. by A. Ronzhin, V. Pshikhopov. Singapore, Springer, 2023, pp. 135–146. DOI: 10.1007/978-981-19-7685-8_9.
20. Gaiduk A.R., Martjanov O.V., Medvedev M.Yu., Pshikhopov V.K., Hamdan N., Farhood A. Neural network based control system for robots group operating in 2-d uncertain environment. *Mekhatronika, Avtomatizatsiya, Upravlenie* = *Mechatronics, Automation, Control*, 2020, vol. 21 (8), pp. 470–479. DOI: 10.17587/mau.21.470-479.
21. Watkins C.J.C.H., Dayan P. Q-learning. *Machine Learning*, 1992, vol. 8, pp. 279–292. DOI: 10.1007/BF00992698.
22. Mnih V., et al. Asynchronous methods for deep reinforcement learning. *Proceedings of Machine Learning Research*, 2016, vol. 48, pp. 1928–1937.
23. Rummery G.A., Niranjan M. *On-line Q-learning using connectionist systems*. Technical Report CUED/F-INFENG/TR 166. Cambridge University Engineering Department, 1994.

24. Sutton R.S., Barto A.G. *Reinforcement learning: an introduction*. 2nd ed. Cambridge, MIT Press, 2018. 526 p.
25. Schulman J., Wolski F., Dhariwal P., Radford A., Klimov O. *Proximal policy optimization algorithms*. arXiv:1707.06347. 2017.
26. Clevert D., Unterthiner T., Hochreiter S. *Fast and accurate deep network learning by exponential linear units (ELUs)*. arXiv:1511.07289. 2015.
27. Premakumar P. *A* (A Star) search for path planning tutorial: MATLAB Central File Exchange*. 2023. Available at: <https://www.mathworks.com/matlabcentral/fileexchange/26248-a-a-star-search-for-path-planning-tutorial> (accessed 28.11.2023).

Для цитирования:

Лукин Д.С., Косенко Е.Ю. Нейросетевой метод планирования пути в двумерном пространстве // Системы анализа и обработки данных. – 2023. – № 4 (92). – С. 55–68. – DOI: 10.17212/2782-2001-2023-4-55-68.

For citation:

Lukin D.S., Kosenko E.Yu. Neurosetevoi metod planirovaniya puti v dvumernom prostranstve [A neural network method for path planning in a two-dimensional space]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2023, no. 4 (92), pp. 55–68. DOI: 10.17212/2782-2001-2023-4-55-68.