

УДК: 004.896

Научная статья

DOI: 10.35330/1991-6639-2024-26-6-208-218

EDN: JSKNGG

Модификация алгоритма глубокого обучения для распределения функций и задач между робототехническим комплексом и человеком в условиях неопределенности и переменности окружающей среды

М. А. Шереужев^{✉1,2}, У Го², В. В. Серебрянный²

¹Московский государственный технологический университет «СТАНКИН»
127055, Россия, Москва, Вадковский пер., 1

²Московский государственный технический университет имени Н. Э. Баумана
105005, Россия, Москва, ул. 2-я Бауманская, 5, корп. 1

Аннотация. В реальном мире условия редко бывают стабильными, что требует от робототехнических комплексов (РТК) способности к адаптации в условиях неопределенности. Синергия человека и робота повышает производительность, однако для этого необходимы эффективные методы распределения задач, учитывающие особенности обеих сторон. Целью работы является определение оптимальных стратегий распределения задач между людьми и РТК и адаптивное управление РТК в условиях неопределенности и изменяющейся среды. Методы исследования. В работе предложен графовый подход к распределению задач, основанный на возможностях человека и робота. В алгоритм обучения с подкреплением встроен механизм памяти LSTM (Long short-term memory) для решения проблемы частичной наблюдаемости, вызванной неточностью измерений сенсоров и шумом окружающей среды. Метод HER (Hindsight Experience Replay) применен для преодоления проблемы скудных вознаграждений. Результаты. Обученная модель продемонстрировала стабильную сходимость, достигая высокого уровня успешности манипуляции объектами. Интеграция методов LSTM и HER в обучение с подкреплением позволяет успешно решать вопросы распределения задач между человеком и роботом в условиях неопределенности и изменяющейся среды. Предложенный метод можно применять в различных сценариях для РТК в сложных и изменяющихся условиях.

Ключевые слова: взаимодействие человека и робота, адаптивный алгоритм управления, распределение задач, обучение с подкреплением

Поступила 13.11.2024, одобрена после рецензирования 06.12.2024, принята к публикации 11.12.2024

Для цитирования. Шереужев М. А., Го У, Серебрянный В. В. Модификация алгоритма глубокого обучения для распределения функций и задач между робототехническим комплексом и человеком в условиях неопределенности и переменности окружающей среды // Известия Кабардино-Балкарского научного центра РАН. 2024. Т. 26. № 6. С. 208–218. DOI: 10.35330/1991-6639-2024-26-6-208-218

Modification of a deep learning algorithm for distributing functions and tasks between a robotic complex and a person in conditions of uncertainty and variability of the environment

M.A. Shereuzhev^{✉1,2}, Wu Guo², V.V. Serebrenny²

¹Moscow State University of Technology STANKIN
127055, Russia, Moscow, build. 1 Vadkovsky street

²Moscow State Technical University named after N. E. Bauman
105005, Russia, Moscow, build. 5 corps 1 Baumanskaya street

Abstract. In the real world, conditions are rarely stable, which requires robotic systems to be able to adapt to uncertainty. Human-robot collaboration increases productivity, but this requires effective task allocation methods that consider the characteristics of both parties. The aim of the work is to determine optimal strategies for distributing tasks between people and collaborative robots and adaptive control of a collaborative robot under uncertainty and a changing environment. Research methods. The paper develops a graph-based approach to task allocation based on the capabilities of a human and a robot. The LSTM memory mechanism is built into the reinforcement learning algorithm to solve the problem of partial observability caused by inaccurate sensor measurements and environmental noise. The Hindsight Experience Replay method is used to overcome the problem of sparse rewards. Results. The trained model demonstrated stable convergence, achieving a high level of success rate of manipulation of objects. The integration of LSTM and HER methods into reinforcement learning allows solving the problems of distributing tasks between a human and a robot under uncertainty and a changing environment. The proposed method can be applied in various scenarios for collaborative robots in complex and changing conditions.

Keywords: human robot interaction, adaptive control algorithm, task distribution, reinforcement learning

Submitted 13.11.2024,

approved after reviewing 06.12.2024,

accepted for publication 11.12.2024

For citation. Shereuzhev M.A., Guo Wu, Serebrenny V.V. Modification of a deep learning algorithm for distributing functions and tasks between a robotic complex and a person in conditions of uncertainty and variability of the environment. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2024. Vol. 26. No. 6. Pp. 208–218. DOI: 10.35330/1991-6639-2024-26-6-208-218

ВВЕДЕНИЕ

В условиях современного тренда на повсеместную цифровизацию и повышение производительности труда за счет внедрения киберфизических систем робототехнические комплексы (РТК) становятся все более важными элементами различных сфер деятельности – от промышленности и медицины до сельского хозяйства и решения специальных задач. Системы, предполагающие взаимодействие РТК и людей в реальном времени, открывают новые горизонты для повышения эффективности труда. Однако несмотря на значительные достижения в области робототехники, многие аспекты распределения функций и задач между РТК и людьми остаются нерешенными, особенно в условиях неопределенности и переменности окружающей среды. В реальном мире условия редко бывают стабильными и предсказуемыми. Традиционные методы планирования и управления, основанные на статических сценариях, не подходят для таких задач, так как они не могут учесть всю возможную вариативность реальных условий [1].

Человек остается незаменимым элементом в системах, где требуется высокая степень гибкости, творчества и принятия решений в условиях неопределенности. РТК в свою очередь обладает способностью выполнять рутинные, опасные и высокоточные задачи с высокой эффективностью и точностью. Синергия между человеком и РТК может значительно повысить общую производительность системы. Однако для достижения этой синергии необходимо разработать эффективные методы распределения задач, которые учитывают сильные и слабые стороны как людей, так и роботов. Например, в задаче сборки промышленной продукции РТК может выполнять точные механические операции, в то время как человек контролирует процесс и принимает решения в случае возникновения нестандартных ситуаций.

Одним из ключевых направлений в области обеспечения взаимодействия человека и РТК является разработка адаптивных алгоритмов управления, которые позволяют РТК изменять свое поведение в зависимости от изменений в окружающей среде. Среди них популярными исследованиями последних лет являются такие алгоритмы, как машинное обучение с учителем и обучение с подкреплением [2–3]. Роботы могут учиться на основе опыта и данных датчиков и становиться более устойчивыми к неопределенности.

Цель данного исследования заключается в определении оптимальных стратегий распределения задач между людьми и РТК и планировании движения РТК в условиях неопределенности и переменности окружающей среды для повышения общей эффективности и надежности выполнения задач.

ОБЩАЯ АРХИТЕКТУРА СИСТЕМЫ

На рис. 1 приведена общая структурная схема системы, в которой используются частично наблюдаемый марковский процесс принятия решений (POMDP) и метод глубокого обучения с подкреплением (Deep Reinforcement Learning, DRL). Компоненты системы: алгоритм DRL, POMDP, интерфейс взаимодействия человек-РТК, источник сенсорных данных и система мониторинга.



Рис. 1. Структурная схема системы управления взаимодействием робота и человека

Fig. 1. Structural diagram of the control system for interaction between a robot and a human

Пользователь взаимодействует с системой через интерфейс взаимодействия человек-робот. Этот интерфейс позволяет пользователю задавать задачи и получать обратную связь от системы. Интерфейс взаимодействия человек-робот обрабатывает команды от пользователя и передает их в систему мониторинга и сенсорных данных. Сенсорные данные и система мониторинга собирают данные из окружающей среды, включая информацию о состоянии роботов и внешних условиях. Эти данные поступают в POMDP для дальнейшего анализа. POMDP анализирует сенсорные данные и оценивает вероятности различных состояний окружающей среды. На основе этой оценки POMDP формирует возможные стратегии действий. Алгоритм DRL используется для адаптации поведения РТК на основе опыта и исторических данных. Этот модуль обучает РТК оптимальным действиям через пробу и ошибку, улучшая их способность адаптироваться к изменениям в окружающей среде. РТК получает команды и стратегии от системы принятия решений и выполняет задачи в реальной среде, взаимодействуя с пользователем через интерфейс.

ФОРМАЛИЗАЦИЯ ЗАДАЧИ ПЛАНИРОВАНИЯ

В промышленной среде методы, основанные на графах, более подходят для представления и декомпозиции сложных задач. Например, наиболее распространенный ориентированный ациклический граф (Directed Acyclic Graph, DAG) [4] позволяет наглядно отображать последовательность и зависимость между задачами, а также поддерживает параллельность выполнения задач, что делает его подходящим для сценариев сотрудничества человека и РТК.

Рассмотрим задачу сборки маленького табурета, которая может быть разложена на 10 подзадач: T1 – подготовка сиденья табурета; T2, T3, T4, T5 – установка ножек табурета А, Б, В, Г на сиденье; T6, T7, T8, T9 – закрепление винтов на ножки А, Б, В, Г; T10 – проверка качества. T6 должно быть выполнено после завершения T1 и T2, а T10 зависит от завершения T6 – T9. Участниками являются оператор и два коллаборативных робота. Для гетерогенных систем необходимо учитывать различия в индивидуальных навыках и ограничениях роботов и людей [5]. Можно распределить подзадачи следующим образом: T2, T3, T6, T7 поручаются человеку; T4 и T8 выполняет робот 1; T5 и T9 выполняет робот 2. Эти подзадачи могут выполняться одновременно. Подготовку T1 и проверку качества T10 также можно поручить человеку на основе его компетенций. Результат можно представить в виде, показанном на рис. 2.

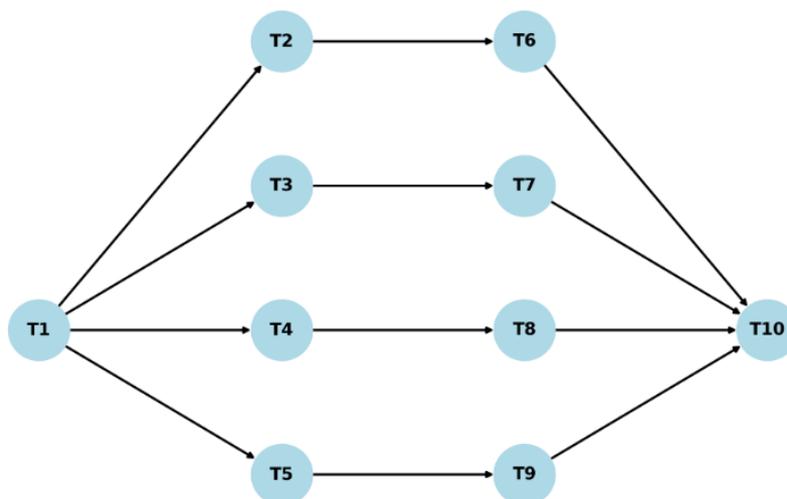


Рис. 2. Пример декомпозиции задачи в виде DAG

Fig. 2. Example of task decomposition in the form of DAG

ФОРМИРОВАНИЕ СТРАТЕГИИ ДЕЙСТВИЙ

DAG обычно предполагает, что процесс выполнения каждой задачи заранее известен и контролируем. Однако выполнение задач в реальном мире часто подвергается воздействию неопределенности и изменениям окружающей среды. Например, измеренные данные датчиков могут содержать шум и неточности. Частично наблюдаемый POMDP обрабатывает такие неопределенности с помощью вероятностной модели. Она позволяет системе динамически корректировать стратегию на основе наблюдений и обновления убеждений в условиях изменяющейся среды, чтобы обеспечить оптимальные решения.

POMDP с дискретным временем определяется 7-кратным кортежем $(S, A, T, R, O, Z, \gamma)$, где S обозначает набор состояний системы, A представляет собой набор действий РТК, $T(s'|s, a)$ обозначает набор условных вероятностей перехода между состояниями, $R(s, a)$ – это функция вознаграждения, $O(z|s', a)$ представляет собой набор условных вероятностей наблюдения, Z обозначает набор возможных наблюдений, $\gamma \in [0,1)$ – это коэффициент дисконтирования.

Поскольку в модели POMDP истинное состояние системы не может быть полностью наблюдаемым, необходимо ввести состояние убежденности по формуле

$$b_t(s) = T(s_t = s | h_t, b_0),$$

которое представляет собой распределение плотности вероятности на пространстве состояний S и отражает предположение или оценку РТК о возможном состоянии системы s на основе текущих наблюдений и истории взаимодействий.

Формула байесовского обновления [6] определяется как

$$b_{t+1}(s') = \eta \cdot O(z|s', a) \cdot \sum_{s \in S} T(s'|s, a) \cdot b_t(s),$$

где η – нормализующая константа, определяемая как

$$\eta = 1 / \sum_{s' \in S} O(z|s', a) \sum_{s \in S} T(s'|s, a) \cdot b_t(s).$$

Задача РТК является выбором оптимальной стратегии $\pi(a | b)$ и максимизацией ожидаемого накопленного дисконтированного вознаграждения по формуле

$$V^\pi(b_0) = E[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) | \pi, b_0].$$

Процесс сборки маленького табурета можно определить как POMDP. Совместное пространство состояний включает информацию о человеке, РТК и непредвиденных обстоятельствах, и каждое состояние s может быть представлено как $s = (s_{r1}, s_{r2}, s_g)$, где s_{r1}, s_{r2} – текущий ход выполнения задачи двух роботов, s_g – состояние объекта, манипулируемого роботом.

Пространство действия РТК A – это декартово смещение концевоего эффектора в трех направлениях. Пространство наблюдений совпадает с пространством состояний. Функция вознаграждения настроена следующим образом: при выполнении конечной задачи присуждается вознаграждение 1, при невыполнении – 0.

АДАПТАЦИЯ ПОВЕДЕНИЯ РОБОТОВ С ПРИМЕНЕНИЕМ DRL

В среде сотрудничества человека и РТК использование DRL может снизить риск столкновений, связанных с присутствием и движением человека. Задачи с непрерывным пространством действий можно обучать с использованием алгоритма DDPG (Deep

Deterministic Policy Gradients) [7], который сочетает в себе преимущества методов актор-критик (Actor-Critic). DDPG использует две нейронные сети: акторную сеть и критическую сеть. Акторная сеть $\mu(s | \theta^\mu)$ принимает состояние s и выдает действие a , а критическая сеть $Q(s, a | \theta^Q)$ оценивает ценность действия a в состоянии s .

Функция потерь для критической сети определяется по формуле

$$L(\theta^Q) = E_{s,a,r,s'}[(Q(s, a | \theta^Q) - y)^2],$$

где целевое значение y вычисляется по формуле $y = r + \gamma Q(s', \mu(s' | \theta^\mu) | \theta^Q)$, и γ – коэффициент дисконтирования.

Обновление акторной сети производится по градиенту политики и определяется по формуле

$$\nabla_{\theta^\mu} J \approx E_s[\nabla_a Q(s, a | \theta^Q)|_{a=\mu(s)} \nabla_{\theta^\mu} \mu(s | \theta^\mu)].$$

В задачах, где данные обладают временной зависимостью, стандартный DDPG может не учитывать информацию из предыдущих состояний. Для решения модели POMDP можно интегрировать в акторную и критическую сети слои LSTM (Long Short-Term Memory) [8], которые способны запоминать долгосрочные зависимости и исторические наблюдения. LSTM описывается по формулам:

$$\begin{aligned} f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), \\ i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i), \\ o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \\ \tilde{C}_t &= \tanh(W_C[h_{t-1}, x_t] + b_C), \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \\ h_t &= o_t \odot \tanh(C_t), \end{aligned}$$

где f_t, i_t, o_t – соответственно сигналы забывания, входа и выхода; C_t – состояние ячейки; h_t – скрытое состояние; σ – сигмоида; \tanh – гиперболический тангенс; \odot – покомпонентное умножение. Структура ячейки LSTM показана на рис. 3.

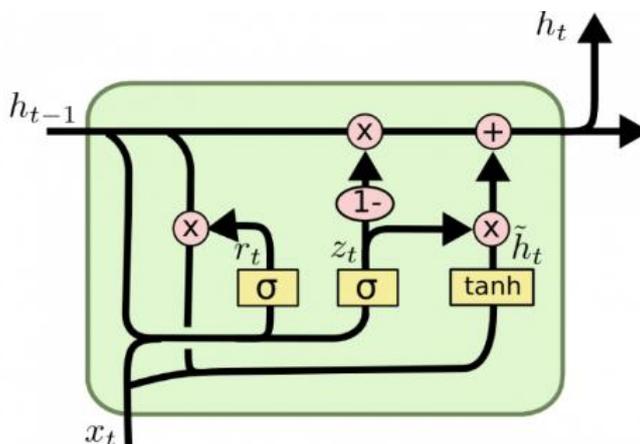


Рис. 3. Структура ячейки LSTM

Fig. 3. LSTM cell structure

Для учета временной информации нужно модифицировать архитектуру DDPG. Акторная сеть теперь зависит не только от текущего состояния, но и от предыдущего скрытого состояния, определяемая как

$$a_t = \mu(s_t, h_{t-1} | \theta^\mu).$$

Критическая сеть оценивает ценность текущего действия с учетом последовательности предыдущих состояний $Q(s_t, a_t, h_{t-1} | \theta^Q)$. Обновления параметров сетей производятся с учетом градиентов по временным последовательностям, что позволяет РТК учитывать контекст при принятии решений. LSTM служит механизмом памяти, который преобразует историю взаимодействий $h_t = \{z_0, a_0, z_1, a_1, \dots, a_{t-1}, z_t\}$, полученную в модели POMDP, в скрытое состояние. Интегрированная с LSTM структура сети DDPG для обучения показана на рис. 4.

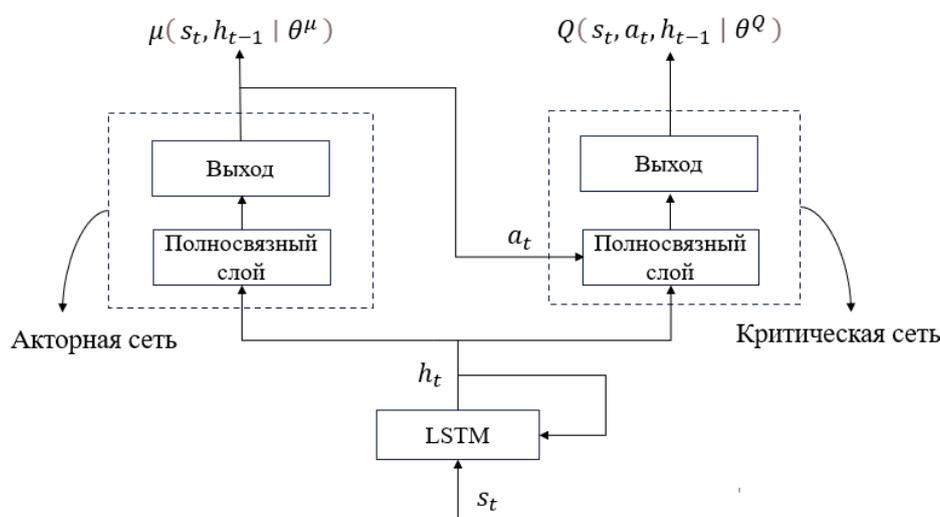


Рис. 4. Структура сети LSTM-DDPG

Fig. 4. LSTM-DDPG network structure

Для задач с чрезмерно скудными вознаграждениями трудно обучить полезную стратегию. Когда положительные вознаграждения скудные, стандартные алгоритмы могут не получать достаточного количества полезных сигналов для обновления политики, что приводит к медленному и неэффективному обучению. Hindsight Experience Replay (HER) [9] – метод, позволяющий агенту учиться из каждого опыта, даже если изначальная цель не была достигнута. Идея HER заключается в том, чтобы переопределять цели после завершения эпизода, используя фактически достигнутые состояния в качестве новых целей. При использовании HER сохраняются переходы в виде кортежей (s, a, r, s', g) , где g – цель. Вознаграждение пересчитывается с учетом новой цели по формуле

$$r = r(s, a, g') = \begin{cases} 0, & \text{если } \phi(s') = g', \\ -1, & \text{иначе,} \end{cases}$$

где $\phi(s')$ – функция, отображающая состояние в достижимую цель. Переопределенные переходы (s, a, r, s', g') используются для обновления политики акторной сети и обучения критической сети.

МЕТОДОЛОГИЯ ПРОВЕДЕНИЯ ЭКСПЕРИМЕНТА И РЕЗУЛЬТАТЫ

Чтобы проверить, может ли предложенная модифицированная структура сети решить проблему POMDP, была выбрана готовая среда 'FetchPickAndPlace-v3', разработанная

OpenAI Gym [10], как показано на рис. 5. Подзадача манипуляции объектами, возложенная на робота, в основном включает захват и перемещение объектов. Цель обучения состоит в том, чтобы научить робота захватывать объект и перемещать его в заданное положение, обозначенное красным шаром. В начале каждого эпизода положение объекта и цели изменяется в небольшом диапазоне.

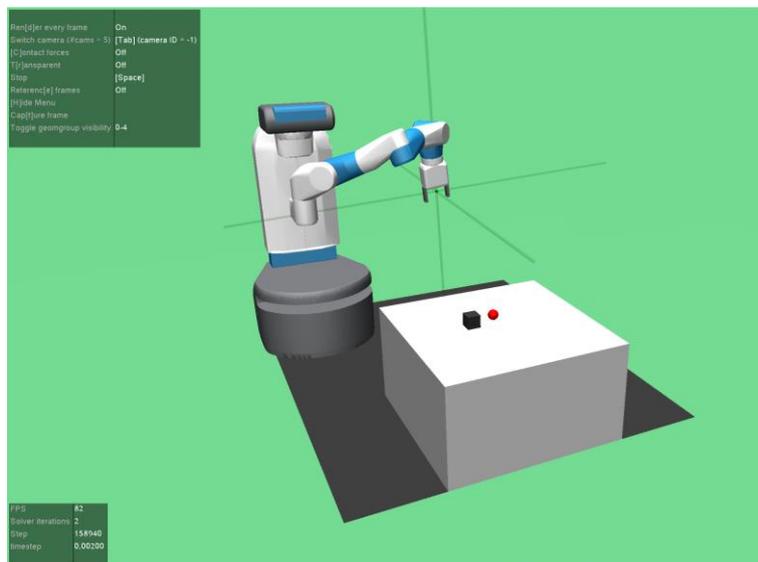


Рис. 5. Готовая среда для обучения

Fig. 5. Ready-made learning environment

Пространство состояний в этой среде состоит из позиции и ориентации концевой эффектора робота, скорости по каждому из измерений, позиции и ориентации объекта и его относительного положения к концевому эффектору, перемещения и скорости захвата. Пространство действия включает в себя команды по перемещению манипулятора в трехмерном пространстве и позиционное смещение каждого пальца захвата. Каждое непрерывное действие принимает значение из $[-1, +1]$.

Для обучения модифицированной модели LSTM-DDPG были использованы гиперпараметры, показанные в табл. 1.

Таблица 1. Проектирование гиперпараметров модели для обучения

Table 1. Designing the hyperparameters of the model for training

Параметры	Значение
Скорость обучения акторной сети	0,001
Скорость обучения критической сети	0,001
Размер пакета	256
Коэффициент дисконтирования	0,95
Параметр мягкого обновления	0,005
Максимальный размер буфера воспроизведения	1000000
Количество переопределений целей	4
Стратегия переопределения целей	'future'
Количество эпизодов обучения	5000
Максимальная длина эпизода	500
Длина последовательности для LSTM	10
Размер скрытого слоя	256

Модель была реализована на платформе машинного обучения Pytorch версии 2.5.1. Программирование осуществлялось на языке Python 3.8.8. Обучение и оценка модели LSTM-DDPG проводились на компьютере с процессором Intel i9-13980HX и графическим процессором NVIDIA GeForce RTX 4060 с памятью 8 ГБ.

Для оценки эффективности модели использовалась метрика «Процент успешных эпизодов». Эта метрика определяется как процент эпизодов из каждых 100, в которых РТК успешно захватил и переместил объект в заданное положение. Формально уровень успеха рассчитывается так:

$$\text{Процент успешных эпизодов} = \left(\frac{\text{Количество успешных эпизодов}}{100} \right) \times 100\%.$$

На рис. 6 представлены результаты обучения модели в виде зависимости уровня успеха от количества эпизодов. Из графика видно, что уровень успеха постепенно увеличивается по мере обучения модели. На начальных этапах обучения успехи редки из-за сложности задачи и скудности вознаграждений. Однако благодаря использованию HER и LSTM модель начинает успешно выполнять задачу после определенного количества эпизодов.

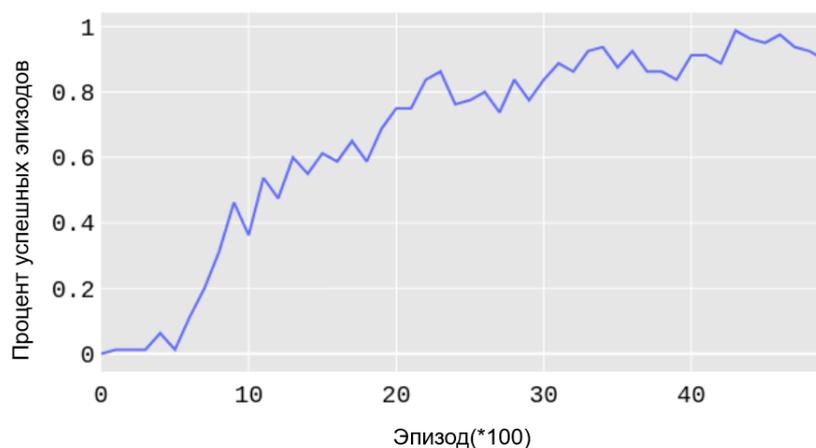


Рис. 6. График зависимости уровня успеха от количества эпизодов

Fig. 6. Graph of success rate versus number of episodes

Модель, обученная с использованием исходного алгоритма DDPG, на протяжении всего процесса обучения показывала нулевой уровень успешности, что свидетельствует о неспособности алгоритма выучить полезную стратегию. После обучения модель с модифицированным алгоритмом продемонстрировала стабильную сходимость, достигая высокого уровня успеха, близкого к 1. Когда объект и цель находятся в любом месте стола, манипулятор может схватить объект и переместить его в заданное место. Использование HER оказалось особенно полезным для задачи манипуляции объектами, поскольку в средах с чрезмерно скудными вознаграждениями сложно выработать полезные стратегии. Интеграция LSTM и HER в архитектуру DDPG показала превосходство над базовым алгоритмом DDPG, особенно в условиях скудных вознаграждений и частичной наблюдаемости среды. Полученные результаты подтверждают эффективность предложенной модифицированной структуры сети в решении задачи POMDP.

ЗАКЛЮЧЕНИЕ

В работе представлен метод распределения задач между человеком и роботом в условиях неопределенности в изменяющейся среде. Задачи разбиваются на основе

графового подхода и распределяются в зависимости от возможностей человека и РТК. Предложено решение, в котором механизм памяти LSTM встроен в алгоритм обучения с подкреплением для решения проблемы частичной наблюдаемости, вызванной неточностью измерений сенсоров и шумом в окружающей среде. Метод HER использован для решения проблемы скудных вознаграждений. Результаты обучения подтверждают эффективность интеграции LSTM и HER в архитектуру DDPG для решения задач POMDP.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Fiore M., Clodic A., Alami R. On planning and task achievement modalities for human-robot collaboration. In *Experimental Robotics: The 14th International Symposium on Experimental Robotics*. Marrakech, Morocco: Springer. 2016. Pp. 293–306.
2. Ghadirzadeh A., Chen X., Yin W. et al. Human-centered collaborative robots with deep reinforcement learning. *IEEE Robotics and Automation Letters*. 2020. Vol. 6(2). Pp. 566–571. DOI: 10.48550/arXiv.2007.01009
3. Qureshi A.H., Nakamura Y., Yoshikawa Y., Ishiguro H. Robot gains social intelligence through multimodal deep reinforcement learning. In *IEEE-RAS. 16th International Conference on Humanoid Robots (humanoids)*. 2016. Pp. 745–751. DOI: 10.48550/arXiv.1702.07492
4. Kwok Y.K., Ahmad I. Static scheduling algorithms for allocating directed task graphs to multiprocessors. *ACM Computing Surveys*. 1999. Vol. 31(4). Pp. 406–471. DOI: 10.1145/344588.344618
5. Malik A.A., Bilberg A. Complexity-based task allocation in human-robot collaborative assembly. *Industrial Robot: International Journal of Robotics Research and Application*. 2019. Vol. 46(4). Pp. 471–480. DOI: 10.1108/IR-11-2018-0231
6. Lucignano L., Cutugno F., Rossi S., Finzi A. A dialogue system for multimodal human-robot interaction. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*. 2013. Pp. 197–204. DOI: 10.1145/2522848.2522873
7. Qiu C., Hu Y., Chen Y., Zeng B. Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications. *IEEE Internet of Things Journal*. 2019. Vol. 6(5). Pp. 8577–8588. DOI: 10.1109/JIOT.2019.2921159
8. Hochreiter S. Long Short-term Memory. Neural Computation MIT-Press. 1997.
9. Andrychowicz M., Wolski F., Ray A. et al. Hindsight experience replay. *Advances in Neural Information Processing Systems*. 2017. Vol. 30.
10. Towers M., Kwiatkowski A., Terry J. et al. Gymnasium: A standard interface for reinforcement learning environments. *arXiv:2407.17032*. 2024. DOI: 10.48550/arXiv.2407.17032

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Финансирование. Исследование выполнено на базе МГТУ «СТАНКИН» при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания (проект № FSFS-2024-0012).

Funding. The study was carried out at MSTU STANKIN with the financial support of the Ministry of Science and Higher Education of the Russian Federation within the framework of the state assignment (project No. FSFS-2024-0012).

Информация об авторах

Шереушев Мадин Артурович, кан. тех. наук, мл. науч. сотр., Центр когнитивных технологий и систем технического зрения, Московский государственный технологический университет «СТАНКИН»;

127055, Россия, Москва, Вадковский пер., 1;

старший преподаватель, кафедра «Робототехнические системы и мехатроника», Московский государственный технический университет им. Н. Э. Баумана;

105005, Россия, Москва, 2-я Бауманская улица, 5, корп. 1;

m.shereuzhev@stankin.ru, ORCID: <http://orcid.org/0000-0003-2352-992X>; SPIN-код: 1734-9056

У Го, аспирант кафедры «Робототехнические системы и мехатроника», Московский государственный технический университет им. Н. Э. Баумана;

105005, Россия, Москва, 2-я Бауманская улица, 5, корп. 1;

ug@student.bmstu.ru, ORCID: <http://orcid.org/0000-0001-8424-4421>; SPIN-код: 9189-9658

Серебряный Владимир Валерьевич, кан. тех. наук, доцент, зав. кафедрой «Робототехнические системы и мехатроника», Московский государственный технический университет им. Н. Э. Баумана;

105005, Россия, Москва, 2-я Бауманская улица, 5, корп. 1;

vsereb@bmstu.ru, ORCID: <http://orcid.org/0000-0003-1182-2117>, SPIN-код: 5410-8433

Information about the authors

Madin A. Shereuzhev, Candidate of Engineering Sciences, Junior Research, Center for Cognitive Technologies and Machine Vision Systems, Moscow State University of Technology STANKIN;

127055, Russia, Moscow, 1 Vadkovsky lane;

Senior Teacher, The Department of Robotic Systems and Mechatronics, Moscow State Technical University named after N.E. Bauman;

105005, Russia, Moscow, 2nd Baumanskaya street 5, bld 1;

m.shereuzhev@stankin.ru, ORCID: <http://orcid.org/0000-0003-2352-992X>; SPIN-code: 1734-9056

Wu Guo, Post-graduate Student at the Department of Robotic Systems and Mechatronics, Moscow State Technical University named after N.E. Bauman;

105005, Russia, Moscow, 2nd Baumanskaya street 5, bld 1;

ug@student.bmstu.ru, ORCID: <http://orcid.org/0000-0001-8424-4421>; SPIN-code: 9189-9658

Vladimir V. Serebrenny, Candidate of Engineering Sciences, Associate Professor, Head of the Department of Robotic Systems and Mechatronics, Moscow State Technical University named after N.E. Bauman;

105005, Russia, Moscow, 2nd Baumanskaya street 5, bld 1;

vsereb@bmstu.ru, ORCID: <http://orcid.org/0000-0003-1182-2117>, SPIN-code: 5410-8433