

ПРЕДОТВРАЩЕНИЕ СТОЛКНОВЕНИЙ БЕЗЭКИПАЖНЫХ СУДОВ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Лама Баракат (Астрахань, Россия – Сирия)

Нами исследуется вопрос о возможности создания методологии применения глубокого обучения с подкреплением к задачам моделирования автономной системы предупреждения столкновений безэкипажных судов, которая обычно способна принимать решения без какого-либо вмешательства человека. Результаты исследования позволяют понять, что интеллектуальная система предупреждения столкновений можно считать одной из основных частей безэкипажного судна.

В работе приводятся ссылки на литературные источники, которые подтверждают или разъясняют представленный материал.

В последние годы достигнут значительный прогресс в развитии морских автономных систем принятия решений и технологий искусственного интеллекта. В результате чего безэкипажное судно (БЭС) стало актуальной темой исследований, проведенных государствами и транснациональными корпорациями. БЭС представляет собой беспилотное интеллектуальное морское транспортное средство, включающее в себя различные датчики и комплекс автономных систем, способных в сложных условиях без какого-либо вмешательства человека принимать решения и выполнять определенные задачи. Среди них, процесс поиска и предотвращения препятствий, который имеет важное значение.

Исследования показывают, что аварии из-за человеческого фактора составляют 89-96% от всех морских столкновений судов [3]. На рисунке 1 иллюстрируется количество морских столкновений судов по годам до 31 марта 2019 года в соответствии со статистикой, представленной японским советом по безопасности на транспорте «Japan Transport Safety Board» [4].

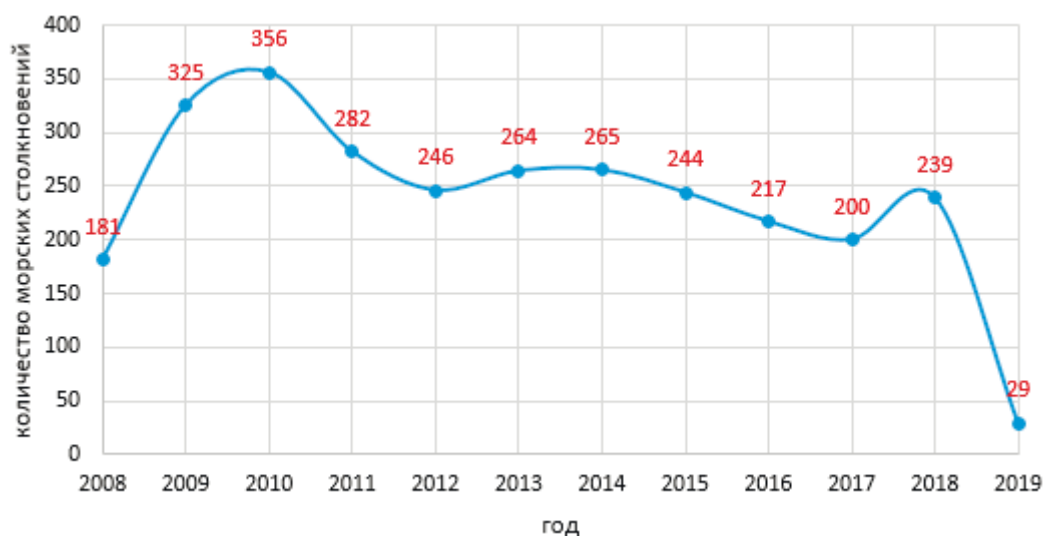


Рис. 1. Статистика морских столкновений судов

На основе представленных на рисунке данных и можно делать выводы о том, что повышение степени автоматизации судна, особенно для системы предупреждения столкновений в соответствии с международными правилами предупреждения столкновений судов (МППСС), можно считать ключевым вопросом для обеспечения безопасности судоходства и исключения человеческих ошибок (сведения их минимум). Исследование автором тематики системы предупреждения столкновений является

основной частью общего исследования в области безэкипажного судоходства, в виду связанности предметных областей.

В ближайшем будущем основная причина морских столкновений (человеческий фактор) будет значительно уменьшена за счет внедрения технологии БЭС и реализации стратегии ИМО по Е-Навигации.

Автономная система предупреждения столкновений БЭС представляет собой алгоритм интеллектуального принятия решений для обеспечения автоматического предотвращения возможных столкновений между судами.

Для понимания сути процесса автономного предотвращения столкновений судов, условно можно его разделить три этапа (рис.2): 1) обнаружение окружающей среды и доступ к степени риска столкновения; 2) установление плана навигации и выполнение управляющей команды для реализации установленного плана и 3) возвращение на свой первоначальный путь, когда риск столкновения исчезает [5].

Следует отметить, что при моделировании системы интеллектуального принятия решений по предупреждению столкновений БЭС необходимо учитывать следующие факторы: 1) переменные динамические характеристики окружающей среды и 2) сложные характеристики движения судна.

Такие проблемы можно решить за счет внедрения и реализации современных технологий искусственного интеллекта (ИИ) в области автоматизации судов. Например, обучение с подкреплением (reinforcement learning) можно считать одним из самых активных направлений исследований в области ИИ. Оно представляет собой разновидность машинного обучения, где обучение происходит через взаимодействие с окружающей средой т.е. обучение с подкреплением является целенаправленным обучением, где агент не учится какие действия предпринять, наоборот, он учится на последствиях своих действий [1]. Глубокое обучение с подкреплением (deep reinforcement learning) – это применение глубокого обучения к обучению с подкреплением.

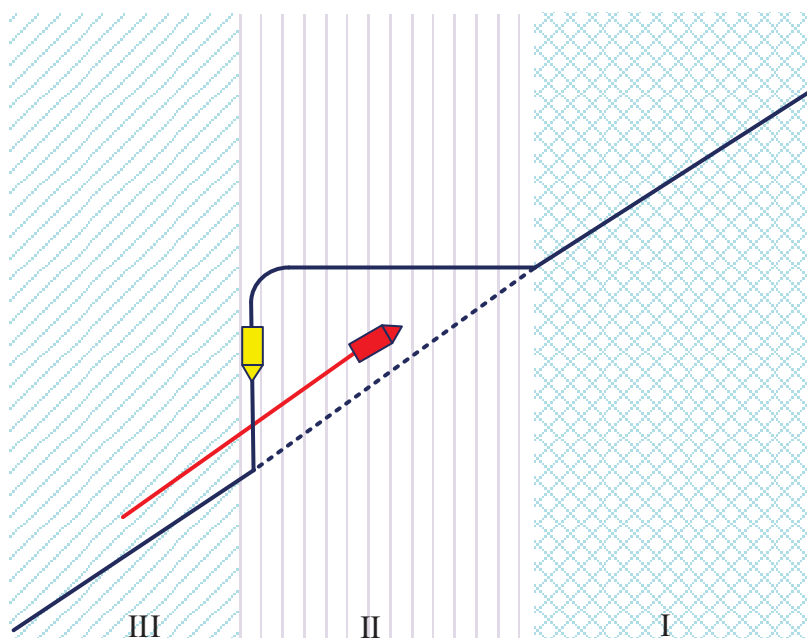


Рис. 2. Поэтапный процесс автономного предотвращения столкновений судов

Моделирование автономной системы предупреждения столкновений БЭС в рамках глубокого обучения с подкреплением обеспечивает оптимальную стратегию управления с помощью глубоких нейронных сетей (deep neural networks) для

аппроксимации функции значения состояния $V(s)$ (state-value function), которая в свою очередь указывает, насколько хорошо для агента (в данном случае агентом является БЭС) находиться в определенном состоянии на временном шаге t .

Взаимодействие БЭС с окружающей средой в рамках обучения с подкреплением изображено на рисунке 3.

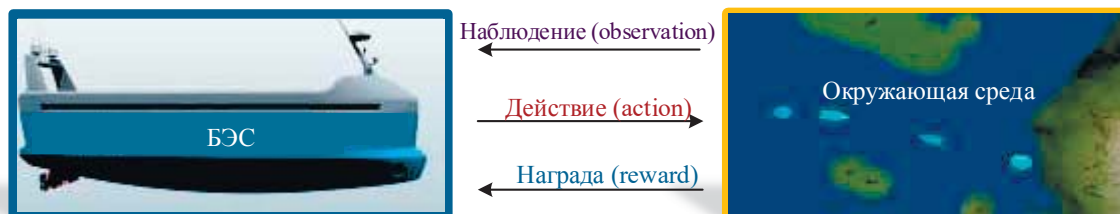


Рис. 3. Взаимодействие БЭС с окружающей средой

Каждую минуту БЭС (агент) может наблюдать полное состояние окружающей среды s_t^t с помощью сенсорной и навигационной систем (радар, АИС, инфракрасная камера, лидар и пр.), чтобы принять решение о действиях, которые оно может предпринять. Судно движется к месту назначения, избегая при этом отображенные препятствия. Информация об окружающей среде являются наградой (от слова reward), данной БЭС окружающей средой, как дополнительную обратную связь о действиях БЭС.

Стоит отметить, что в данном исследованном случае действия БЭС влияют не только на немедленные награды, но и на последующие состояния, и через эти будущие награды. При этом можно использовать марковский процесс принятия решений (Markov decision process), который можно описать как случайный процесс управления с дискретным временем, для описания математической идеализированной формы задачи обучения с подкреплением [2].

Марковский процесс принятия решений определяется следующими пятью элементами (S, A, P, r, γ) , и на момент t БЭС находится в состоянии $s_t \in S$ (где S – множество состояний) и выберет действие $a_t \in A$ (где A – множество действий). В результате чего БЭС приходит в новое состояние s_{t+1} и соответствующее вознаграждение $r(s_t, a_t)$, которое отдается БЭС с новым состоянием s_{t+1} , рандомизированным в соответствии с функцией перехода между состояниями $P(s_t, s_{t+1}) = P(s_{t+1} | s_t, a_t)$ [6]. Поэтому целью агента является максимизация совокупного вознаграждения, которое определяется по следующей формуле [1]:

$$R_t = \sum_{t=0}^n \gamma^t r_{t+1} \quad (1)$$

где $\gamma \in [0, 1]$ – коэффициент дисконтирования.

Кроме того, стратегия (policy) π – функция, которая для данного состояния s выдает распределение вероятностей на множестве действий A [7]. В данном случае стратегия БЭС (агента) можно рассматривать в качестве контроллера судна, который в свою очередь дает исполнительным механизмом (рулю судна, двигателю и пр.) необходимые команды для управления судном и избегания возможных препятствий. Оптимальная стратегия марковского процесса принятия решений π^* определяется в зависимости от глубокого Q-обучения (deep Q-learning). Основная идея подхода

глубокого Q-обучения заключается в запоминании исторической информации, которую его алгоритм выполняет в этой среде [8].

Таким образом, в настоящее время проводят много исследований, связанных с безэкипажным судоходством, которое в ближайшем будущем может принести пользу морскому обществу во многих отношениях, таких как уменьшение воздействия на окружающую среду и сокращение числа столкновений за счет разработки систем принятия решений, например, в рамках глубокого обучения с подкреплением, которое может решить проблему автоматического обучения оптимальных решений с течением времени.

Данная тема имеет перспективы развития не только в области морского судоходства, но и в области беспилотных автомобилей и робототехники.

Литература

1. **Ravichandiran S.** Hands-On Reinforcement Learning with Python: Master Reinforcement and Deep Reinforcement Learning Using OpenAI Gym and TensorFlow. Birmingham: Packt Publishing Ltd., 2018. P.764.
2. **Sutton R.S., Barto A.G.** Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning series). Cambridge: The MIT Press, 2018. p.548.
3. **Rothblum A. M.** Human error and marine safety // National Safety Council Congress and Expo, Orlando, FL. 2000. P.1-10.
4. Japan Transport Safety Board. [Электронный ресурс]. URL: http://www.mlit.go.jp/jtsb/statistics_mar.html. (дата обращения: 07.06.2019).
5. **You Y.J. et al.** A study on the collision avoidance system of a ship considering the effects of speed dependent coefficients // The Twentieth International Offshore and Polar Engineering Conference. International Society of Offshore and Polar Engineers, 2010. pp.401-408.
6. **Levy K.Y., Shimkin N.** Unified inter and intra options learning using policy gradient methods // European Workshop on Reinforcement Learning. Springer, Berlin, Heidelberg, 2011. pp. 153-164.
7. **Николенко С., Кадурин А., Архангельская Е.** Глубокое обучение. СПб.: Питер, 2018. 480 с.
8. **Wang C. et al.** Research on intelligent collision avoidance decision-making of unmanned ship in unknown environments // Evolving Systems. 2018. Pp. 1-10.