

**ПРИМЕНЕНИЕ МЕТОДОВ ОБУЧЕНИЯ С
ПОДКРЕПЛЕНИЕМ К ЗАДАЧЕ СТАБИЛИЗАЦИИ
СИСТЕМЫ «ХИЩНИК-ЖЕРТВА»**

Ларионов Александр Владимирович

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: avlariонов98@gmail.com

Научный руководитель — Смирнов Алексей Игоревич

Задача стабильного и сбалансированного взаимодействия популяций различных видов становится всё более важной с каждым годом, когда климатические изменения и результаты человеческой деятельности привели к тому, что многие виды находятся на грани вымирания [5]. В данной работе рассматривается задача стабилизации динамической системы для модели «хищник-жертва» с помощью методов обучения с подкреплением [1].

Рассматривается модель, описывающая взаимодействие популяций канадской рыси и американского зайца-беляка [2–3]:

$$\begin{cases} \dot{H} = rH(1 - \frac{H}{k}) - \frac{aHL}{H + c}, \\ \dot{L} = b\frac{aHL}{H + c} - dL, \end{cases} \quad (1)$$

где H — численность популяции зайцев («жертв»), L — численность популяции рысей («хищников»), r — характеристика естественного прироста в популяции зайцев, k — ёмкость среды (то есть какую максимальную численность популяции зайцев среда способна прокормить), a — параметр взаимодействия между видами, показывающий, как количество зайцев уменьшается в зависимости от количества рысей, c — коэффициент, отвечающий за скорость потребления в случае небольшой численности «жертв», b — характеристика роста популяции рысей, d — уровень смертности рысей [1–2].

Считаем, что мы можем управлять рождаемостью зайцев, создавая для них более или менее благоприятные условия [2]. Тогда первое уравнение системы (1) примет вид:

$$\dot{H} = (r + u)H(1 - \frac{H}{k}) - \frac{aHL}{H + c}, \quad (2)$$

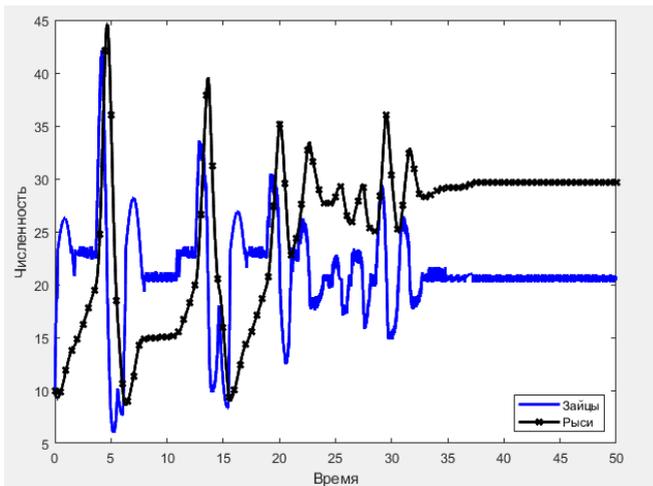
где $u \in [-1, 1]$ - управление.

В качестве основного метода выбран алгоритм SARSA, так же известный как метод на основе временных различий с интегрированной оценкой ценности стратегии. Методы на основе временных различий сочетают в себе особенности методов Монте-Карло (обучение непосредственно на полученном опыте) и динамического программирования (способность к самонастройке), пользуясь их достоинствами и сглаживая недостатки [4].

Работа алгоритма осуществляется по эпизодам. Вводится дискретное множество состояний, каждое из которых определяется некоторым отрезком значений по каждой из переменных. Также вводится терминальное состояние, которому соответствуют такие значения переменных, при которых становится ясно, что стабилизация не удалась. После попадания в такое состояние происходит повторная инициализация системы начальными значениями и начинается новый эпизод. Если эпизод не заканчивается в течение заданного количества временных шагов, то делаем вывод, что стабилизация удалась.

Поведение системы под воздействием управления, полученного на основе алгоритма SARSA, было смоделировано с помощью среды MatLab. Были построены графики, демонстрирующие изменение численности популяций зайцев и рысей во времени.

Иллюстрации



Динамика численности популяций

Литература

1. Ризниченко Г. Ю., Рубин А. Б. Математические методы биологических продукционных процессов. М.: Издательство Московского университета, 1993.
2. Astrom K. J., Murray R. M. Feedback Systems: An Introduction for Scientists and Engineers. Princeton University Press, 2008
3. MacLulich D. A. Fluctuations in the Numbers of the Varying Hare (*Lepus Americanus*). University of Toronto Press, 1937
4. Sutton R. S., Barto A. G. Reinforcement Learning: An Introduction. Second edition. The MIT Press, 2018
5. Страница о голоценовом вымирании:
https://ru.wikipedia.org/wiki/Holocene_extinction