



DOI: 10.22363/2312-8313-2025-12-3-366-374

EDN: BRCVKY

Research article / Научная статья

Enhancing governmental policy-making in demographics and migration through multi-agent Deep reinforcement learning: a case study with the MADDPG algorithm

Anton V. Dozhdikov  

Institute of Social and Political Studies, FNISSC RAS, Moscow, Russian Federation

 antondnn@yandex.ru

Abstract. The study identifies the main social, political and economic risks associated with the “overproduction” of the elite, the reduction of the middle class, taking into account uncontrolled migration. To mitigate the risks, a general theoretical approach is proposed to optimize the “hyperparameters” of public administration procedures, “upgrade” the decision-making model using hybrid systems based on machine learning. The experiment was conducted for 7 regions with initially random features (the number of regions can be any). During the experiment with the MADDPG algorithm, the author shows the possibility of implementing a balanced migration, socio-economic and resource policy for an arbitrary number of regions in conditions of instability, chaotic, noise processes and interregional migration for an unlimited period while maintaining the main environmental parameters. Trained AI algorithms in joint activities showed population growth, economic growth and development of territories, rational use of available resources (without their depletion), balanced interregional migration. Further direction of the research involves the inclusion of the external migration factor and detailing the factors of interregional migration, economic growth and resource consumption in the context of the social structure of society. The prospect of application are hybrid human-machine control and decision support systems for the sphere of public political administration.

Keywords: demographic structure, labor mobility patterns, overproduction of skilled professionals, sequential decision-making algorithms, neural function approximation, collaborative autonomous agents, state governance frameworks

Acknowledgments. The author expresses gratitude to the teaching staff of the OTUS course on Reinforcement Learning for the knowledge and skills acquired.

Conflicts of interest. The author declares no conflicts of interest.

Article history:

The article was submitted on 10.04.2025. The article was accepted on 30.05.2025.

© Dozhdikov A.V., 2025



This work is licensed under a Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

For citation:

Dozhdikov AV. Enhancing governmental policy-making in demographics and migration through multi-agent Deep reinforcement learning: a case study with the MADDPG algorithm. *RUDN Journal of Public Administration*. 2025;12(3):366–374. <https://doi.org/10.22363/2312-8313-2025-12-3-366-374>

Совершенствование процедур государственного управления в сфере демографии и миграции с помощью мультиагентного Deep reinforcement learning на примере алгоритма MADDPG

А.В. Дождиков  

Институт социально-политических исследований ФНИСЦ РАН, Москва, Россия

 antondnn@yandex.ru

Аннотация. Определены основные социальные, политические и экономические риски, связанные с «перепроизводством» элиты, сокращением среднего класса с учетом неконтролируемой миграции. Для нивелирования рисков предложен общий теоретический подход оптимизации «гиперпараметров» процедур государственного управления, «апгрейда» модели принятия управленческих решений с помощью гибридных систем, основанных на машинном обучении. Проведен эксперимент для 7 регионов с изначально случайными признаками (число регионов может быть любым). В ходе эксперимента с алгоритмом MADDPG показана возможность реализации сбалансированной миграционной, социально-экономической и ресурсной политики для произвольного числа регионов в условиях нестабильности, хаотических, шумовых процессов и межрегиональной миграции на неограниченный период времени при сохранении основных параметров среды. Обученные ИИ-алгоритмы в совместной деятельности показали рост численности населения, экономический рост и развитие территорий, рациональное использование имеющихся ресурсов (без их истощения), сбалансированную межрегиональную миграцию. Дальнейшее направление исследования предполагает подключение фактора внешней миграции и детализацию факторов межрегиональной миграции, экономического роста и потребления ресурсов в разрезе социальной структуры общества. Перспектива применения — гибридные человеко-машинные системы управления и поддержки принятия решений для сферы государственного управления.

Ключевые слова: динамика населения, миграционные потоки, избыток элиты, обучение с подкреплением, глубокое обучение, многоагентные системы, государственная политика

Благодарности. Автор выражает благодарность коллективу преподавателей OTUS курса Reinforcement Learning за полученные знания и навыки.

Заявление о конфликте интересов. Автор заявляет об отсутствии конфликта интересов.

История статьи:

Поступила в редакцию 10.04.2025; принята к публикации 30.05.2025.

Для цитирования:

Дождиков А.В. Совершенствование процедур государственного управления в сфере демографии и миграции с помощью мультиагентного Deep reinforcement learning на примере алгоритма MADDPG // Вестник Российского университета дружбы народов. Серия: Государственное и муниципальное управление. 2025. Т. 12. № 3. С. 366–374. <https://doi.org/10.22363/2312-8313-2025-12-3-366-374>

Introduction

Demographic and migration processes carry risks of instability due to structural imbalances [1]. The classical approach [2] assumes that shocks occur in countries with a large number of young people. But it is a mistake to assume that in the absence of a “youth hillock” [3], which is the cause of the “Arab spring” [4] and the indicator of demographic reproduction of the population in the range of 1.4-1.5, there are no prerequisites for crises: The reduction of the population leads to a decrease in the economic potential and the “division” of the remaining. An analogue of the “youth hill” takes place in the North Caucasus Federal District [5], as well as in the countries of Central Asia, which have long been donors of migration [6].

The number of the elite can grow¹ even in conditions of economic recession and population decline. The result is its “overproduction” [7], and crises if there are no areas for expansion, similar to the eras of the Crusades and Great Geographical Discoveries².

The upper segment of the middle class is poorly exposed to incentive measures. The middle segment of the middle class is sensitive to government support to certain limits. The lower middle class segment is very sensitive, and government support measures can improve the situation, as well as for the lower middle class segment. In relation to groups of chronic poverty, the precariat [8] and marginal strata, dependence needs to be studied.

To implement demographic policy, the government applies economic measures (maternity capital, preferential seed mortgage). Similarly, non-material ones are the promotion of the ideas of natalism [9] in the context of traditional values [10], including a ban on the propaganda of “anti-natalist” ideologies, suggesting that they have an impact on the matrimonial strategies of young people [11].

By today, direct incentive measures have become ineffective: the preferential mortgage program is the reason for the rise in real estate prices, as a result of which the average area of new apartments has decreased by an average of one room³. An apartment without a mortgage is not available

¹ According to a 2022 study, the number of children of the spouses of Russian governors averaged 2.27. Quoted by: Kolebakina-Usmanova E. Minchenko Consulting on the wives of governors. *Business Newspaper*: 01.07.2022. URL: <https://www.business-gazeta.ru/article/552026> (accessed: 08.03.2025). (In Russ.).

² Dozhdikov AV. “Space Conquistadors” by Elon Musk. *Plas*. 06.07.2024. URL: <https://plusworld.ru/journal/2024/plas-5-313-2024/kosmicheskie-konkistadory-ilona-maski/> (accessed: 08.03.2025). (In Russ.).

³ The average area of apartments in Moscow has decreased by a quarter in 20 years. *RBC. Realty*. 04.10.2022. URL: <https://realty.rbc.ru/news/615ae2589a79477d48835e97> (accessed: 08.03.2025). (In Russ.).

for the middle and lower middle segments. The upper segment of society, which can pay a mortgage and receives benefits, buys investment apartments, maintaining high prices. As a result, the price of an apartment (using Moscow as an example) increased by 14.66 times from February 2000 (19,743 rubles) to February 2025 (274,856 rubles)⁴, outpacing inflation. The increase in prices (deterioration of housing affordability) and the decrease in living area negatively affect birth rate.

The second problem is the expensive “social reproduction” for the middle class, not only in terms of finances, but also the time spent by parents on maintaining social status and then passing it on to their children. Therefore, the methods of natalist propaganda do not work: “Wealthier married couples lose more compared to less well-off spouses, because their time is valued more”⁵.

The third problem is uncontrolled migration processes: the “new citizens” are poorly integrated into society, preferring enclaves (future ghettos), monopolizing certain sectors of the economy. However, due to the principle of equality of citizens before the law, they use housing certificates and other benefits to a greater extent than the “old citizens”.

Two “youth humps” are forming in Russia — among the elite and among the “new citizens”. The middle class, which cements society, is shrinking both in relative and absolute terms, being squeezed from several sides.

In the case of an “archaization” scenario due to the proliferation of “economic enclaves” on the one hand, and a strict conservative, traditionalist government policy, there is a risk of conflicts both on national and religious grounds and on socio-economic grounds.

An alternative to the “archaization” scenario is effective management, dynamic maneuvering of social benefits, incentive measures, cultural policy, and migration processes. This approach involves broadcasting the experience of some territories to others and flexible coordination from the federal center acting as a “critic”.

The introduction of such a technique is associated with a theoretical approach to the political system as a complex ensemble model of management and political decision-making [12]. Adjusting its hyperparameters or “upgrading” significantly improves performance and quality metrics.

⁴ *Housing cost index in Moscow*. URL: <https://www.irm.ru/gd/s-2000-goda/> (accessed: 08.03.2025). (In Russ.).

⁵ HSE experts: due to rising incomes, the likelihood of having children in the family is decreasing. *Kommersant. Economy*. 22.12.2023. URL: <https://www.kommersant.ru/doc/6414088> (accessed: 08.03.2025). (In Russ.).

To implement public policy, it is necessary to use artificial intelligence capable of self-learning, the scientific novelty of the research lies in the application of the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) method, used for cyber-physical systems and UAV swarm control for the needs of public policy management.

The purpose of the study is to build a model of interaction between agents managing regional development based on a set of indicators, to illustrate the possibility of using AI tools for demographic growth, stabilization of resource costs, and migration processes.

Materials and methods

Complex processes require the use of deep reinforcement learning techniques. The algorithm is used for multi-agent dynamic environments. It allows agents to learn how to interact with each other: each learns the optimal strategy based on their own observations and the actions of others, which makes the method adaptive and flexible.

The main problem of reinforcement learning in DQN-type models (Deep Q-Learning, [13]) is the “non-statisticity” of the environment, the policy gradient suffers from an increase in variance with an increase in the number of agents. A once-trained algorithm is difficult to scale (there are 89 regions in the Russian Federation). The model created for individual regions or the developed policy is unsuitable for the rest.

The MADDPG algorithm [14] is used in the control of unmanned [15] and military systems [16], is effective for mixed cooperative and competitive environments [17], and is one of the prototypes of a promising “collective artificial intelligence” [18; 19]. MADDPG extends the Deep Deterministic Gradients (DDPG) method for cooperative or competitive work in complex environments. MADDPG⁶ takes into account the policies of other agents and is able to successfully learn policies that require complex coordination between AI agents. A training regime is introduced using an ensemble of policies for each agent to create sustainable multi-agent communities.

The method assumes centralized planning and decentralized execution: each agent has direct access to local observations, during training all agents are controlled by a central module, during testing it is deactivated, and agents with their own policies and local data remain. MADDOG includes a minimax

⁶ Lowe R. *Multi-agent actor-critic for mixed cooperative-competitive environments*: arXiv preprint; 2020. 9 p. URL: <https://arxiv.org/abs/1706.02275> (accessed: 07.03.2025).

strategy that regulates the policy of each agent so that it acts optimally even in the worst case.

For management, it is important to be able to implement multi-agent collaboration and collaboration, set common goals, share information, collaborate, and synchronize actions — which is much more effective than conventional reinforcement learning methods in complex environments [20].

Results

PyTorch was used for the experiment, and a dynamic model of the environment was created without detailing individual socio-economic indicators: the population depends on the attractiveness of policies implemented in cities; economic growth is related to the population and current policies; resources are depleted in proportion to economic activity.

The function of rewarding agents includes: rewards for maintaining population balance, economic growth, resource sustainability, and coordination with neighbors. Agents learn to balance between attracting the population and conserving resources; clusters of regions with coordinated policies are formed: the system comes to a state of dynamic equilibrium with moderate fluctuations in indicators due to random factors.

7 profiles of regions were generated (the number can be any). Training was carried out over 100 epochs with 200 steps each, since the task requires maintaining balance for a long time. Pseudorandom indicators are fixed at the start, and the results of the experiment can be reproduced. In the code⁷ there is the environment model that includes the creation of random profiles of regions (with populations from 0.5 to 1.5), a mechanism for calculating migration flows, factors of economic growth and resource consumption, and a reward function. The actor-critic architecture has been created, suggesting a decentralized policy for each decision-making agent and a centralized value function for the critic. The MADDPG class has been implemented, which has the function of adding chaotic effects and creating agents based on the number of regions.

After episode 20, the AI system learned (the total reward is approximately 100) and was able to maintain this state for the remaining epochs. The graph (Fig.) shows the reward function and a relatively stable “plateau” of learning after episode 20.

⁷ The code and results are publicly available. URL: https://github.com/AntonDozhnikov/AntonDozhnikov/blob/main/Demography_MADDPG.ipynb (accessed: 08.03.2025).

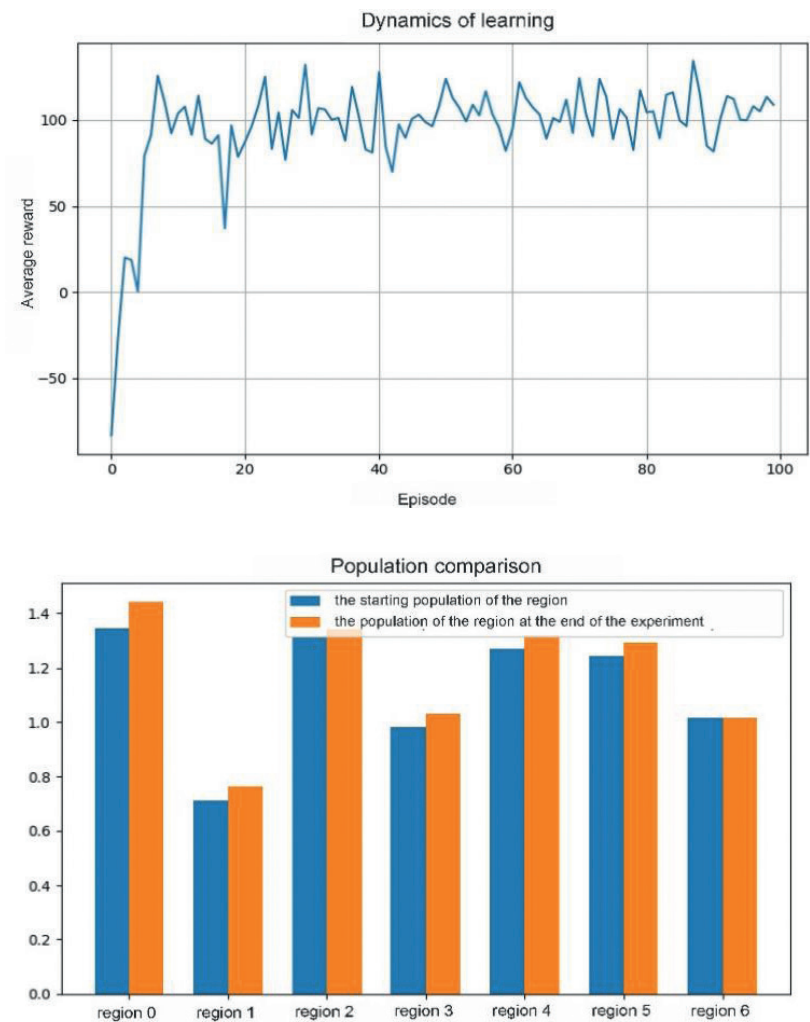


Illustration of the training dynamics of AI agents, along with the initial and final population levels across regions

Source: Implemented by A.V. Dozhdikov in the Python environment using the PyTorch deep learning framework for the realization of the MADDPG agent/class.

The results of the experiment showed an overall increase in the population, while maintaining steady economic growth, balancing migration processes, and avoiding resource depletion.

Conclusion

The basic limitation of the experiment is the training time and the performance of the equipment, including graphics accelerators. The strategy for overcoming limitations involves parallelizing learning, or changing

the architecture, for example, 2-3-stage training: first, several “innovation agents” are trained, then untrained ones are added to them, with whom the first share their experience. The second option involves dividing regions into groups of similarity based on socio-economic characteristics using standard clustering methods with training of representatives and subsequent dissemination of experience.

The AI agent of regional management, based on its own experience and the experience of neighbors, will be able to: predict migration trends, develop initiatives to improve demographic indicators; model scenarios aimed at improving living conditions; adapt migration strategy; optimize the supply of resources and the service sector; flexibly respond to changes in demand for services and plan infrastructure development; organize interact with other AIs by sharing data and experiences.

The MADDPG method provides a toolkit for managing social processes. The integration of multi-agent systems into management allows government agencies to respond more flexibly and effectively to changes in demography and migration. Such a class of methods can become a key element in management, creating sustainable and adaptive systems. The implementation of the “actor-critic” model in MADDPG corresponds to the structure of political governance in the Russian Federation: “critic” is the federal center, “actors” are management agents in districts and regions, which creates the possibility of “upgrading” management and decision-making procedures.

REFERENCES

1. Zinkina YuV, Shulgin SG. “Youth bulge” as a factor of sociopolitical instability. *Bulletin of Moscow University. Series 27: Global Studies and Geopolitics*. 2020;(1):41–52. (In Russ.). <https://doi.org/10.56429/2414-4894-2020-31-1-41-52> EDN: DREPPQ
2. Goldstone JA. *Revolution and rebellion in the early modern world*. Berkeley: University of California Press; 1991.
3. Nefedov SA. “Youth bulge” and the first Russian revolution. *Sociological Research*. 2015;(7):140–147. (In Russ.). EDN: UCFOCB
4. Korotaev AV, Isaev LM. The bumps and faults revolution. *Ehkspert*. 2012;(30/31):Special Issue:7–10. (In Russ.).
5. Murtuzalieva DD, Simagin YuA, Vankina IN. Dynamics of the population of the North Caucasian regions of Russia in 2010–2022. *Population*. 2022;25(3):33–45. (In Russ.). <https://doi.org/10.19181/population.2022.25.3.3> EDN: BSBPER
6. Akramov ShYu, Blinichkina NYu. Demographic security in the context of international migration. *DEMIS. Demographic Research*. 2023;3(2):28–39. (In Russ.). <https://doi.org/10.19181/demis.2023.3.2.2> EDN: CESKQP
7. Turchin P. Political instability may be a contributor in the coming decade. *Nature*. 2010;463:608. <https://doi.org/10.1038/463608a>
8. Popov AV. From precarious employment to the precariat. *Sociological Research*. 2020;(6):155–160. (In Russ.). <https://doi.org/10.31857/S013216250009300-3> EDN: YOXAYW

9. Klupt MA. Family and fertility issues in value conflicts during the 2010s. *Sociological Research*. 2021;(5):36–46. (In Russ.). <https://doi.org/10.31857/S013216250014119-3> EDN: TAKIRG
10. Svadbina TV, Nemova OA. The Russian family as a guardian and translator of traditional national values. *Vestnik of Minin University*. 2023;11(4):14. (In Russ.). <https://doi.org/10.26795/2307-1281-2023-11-4-14> EDN: EHHNCH
11. Blagorozheva ZhO, Shapovalova IS. The influence of alternative values and attitudes on the matrimonial strategies of youth. *Sotsial'naya politika i sotsiologiya*. 2024;23(2):30–39. (In Russ.). <https://doi.org/10.17922/2071-3665-2024-23-2-30-39> EDN: HTXQXM
12. Dozhdikov AV. Political system as a machine-learning model. *Technologies of Social and Humanitarian Research*. 2024;(2):9–24. (In Russ.). EDN: MTPDDQ
13. Li Sh. *Reinforcement learning for sequential decision and optimal control*. 1st ed. Singapore: Springer; 2023. <https://doi.org/10.1007/978-981-19-7784-8>
14. Fu X, Wang H, Xu Z. Cooperative pursuit strategy for multi-UAVs based on DE-MADDPG algorithm. *Acta Aeronautica et Astronautica Sinica*. 2022;43(5):325311. <https://doi.org/10.7527/S1000-6893.2021.25311> EDN: XBKXBXQ
15. Liu Bo, Wang Sh, Li Q, Zhao X, Pan Yu, Wang Ch. Task assignment of UAV swarms based on deep reinforcement learning. *Drones*. 2023;7(5):297. <https://doi.org/10.3390/drones7050297> EDN: STKSJG
16. Li W, Chen X, Yu W, Xie M. Multiple unmanned aerial vehicle coordinated strikes against ground targets based on an improved multi-agent deep deterministic policy gradient algorithm. *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*. 2024. <https://doi.org/10.1177/09596518241291185> EDN: WXEFO
17. Wei X, Huang X, Yang LF. et al. Hierarchical RNNs-based transformers MADDPG for mixed cooperative-competitive environments. *Journal of Intelligent and Fuzzy Systems*. 2022;43(1):1011–1022. <https://doi.org/10.3233/JIFS-212795> EDN: HLEHUN
18. Wang Zh, Guo Ya, Li N. Hu Sh, Wang M. Autonomous collaborative combat strategy of unmanned system group in continuous dynamic environment based on PD-MADDPG. *Computer Communications*. 2023;200:182–204. <https://doi.org/10.1016/j.comcom.2023.01.009> EDN: ROHQW
19. Zhao M, Wang G, Fu Q, et al. MW-MADDPG: A meta-learning based decision-making method for collaborative UAV swarm. *Frontiers in Neurorobotics*. 2023;17. <https://doi.org/10.3389/fnbot.2023.1243174> EDN: NPYWPK
20. Chen Zh. DQN-MADDPG Coordinating the multi-agent cooperation. *Highlights in Science, Engineering and Technology*. 2023;39:1141–1145. <https://doi.org/10.54097/hset.v39i.6720> EDN: XKQISV

Information about the author:

Anton V. Dozhdikov — Candidate of Political Sciences, Senior Researcher, UNESCO Department, Institute of Social and Political Studies, FNISSC RAS, 6 Fotievoy st., bldg. 1, Moscow, 119333, Russian Federation (ORCID: 0000-0002-1069-1648) (SPIN-code: 2208-1891) (e-mail: antondnn@yandex.ru).