Calibration of Agent-Based Model Using Reinforcement Learning

Bing Song
The State Key Laboratory for Management and Control of Complex Systems
Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence
University of Chinese Academy of Sciences
Beijing, China
songbing2020@ia.ac.cn

Gang Xiong
The State Key Laboratory for Management and Control of Complex Systems
Institute of Automation, Chinese Academy of Sciences
The Guangdong Engineering Research Center of 3D Printing and Intelligent Manufacturing, The Cloud Computing Center, Chinese Academy of Sciences
Beijing, China
gang.xiong@ia.ac.cn

Songmin Yu
Fraunhofer Institute for Systems and Innovation Research
Breslauer ,Germany
songmin.yu@isi.fraunhofer.de

Peijun Ye (Correspondence Author)
The State Key Laboratory for Management and Control of Complex Systems
Institute of Automation, Chinese Academy of Sciences
Beijing, China
peijun.ye@ia.ac.cn

Xisong Dong
The State Key Laboratory for Management and Control of Complex Systems
Institute of Automation, Chinese Academy of Sciences
Beijing, China
xisong.dong@ia.ac.cn

Yisheng Lv
The State Key Laboratory for Management and Control of Complex Systems
Institute of Automation, Chinese Academy of Sciences
Beijing, China
yisheng.lv@ia.ac.cn

2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI) | 978-1-6654-3337-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/DTPI52967.2021.9540180

***Abstract*—In the research and application of Agent-Based Models (ABM), parameter calibration is an important content. Based on the existing state transfer equations that link the micro-parameters and macro-states of the multi-agent system, this paper further proposes to introduce Reinforcement Learning when calibrating the parameters. The state transfer of the agent after learning is used to calibrate the micro-parameters of ABM, and the interaction between each agent and multiple other agents is expressed as the parameters of the agent. The application case study of population migration demonstrates that our method can achieve high accuracy and low computational complexity.**

Keywords—Agent-Based Model, Reinforcement Learning, calibration

I. Introduction

Agent-Based Model (ABM) can simulate the decision-making process and behavior patterns of individuals or groups, and can simulate the complex interactions between different individuals and between individuals and the environment. This makes ABM a powerful tool for studying social ecosystems, especially in urban traffic analysis [1,2], population migration [3,4], computational sociology [5], and group behavior analysis [6].

Parameter calibration of ABM refers to calibrating the microscopic parameters of ABM according to the output of the system and the actual reality, so that the simulation of the system can be closer to the actual case. Each agent in ABM is heterogeneous, that is, the interaction between different agents and the environment and other agents is different. The agent also has dynamic characteristics, which greatly increases the complexity of the system. It is difficult to directly obtain specific microscopic parameters in the system from the macroscopic observation results, which makes the calibration of parameters one of the difficulties of ABM.

For the difficulty of calibration, there are currently several types of methods. Nicholas Magliocca *et al.* used the method of global sensitivity analysis to combine variance-based and density-based sensitivity analysis to better understand the impact of parameter values on the output [7]. Fagiolo G. *et al.* used indirect inference methods to infer the model's performance through simulation parameters, which will have better performance when the parameters are not many or the scale is small [8]. Simone Alfarano *et al.* proposed a very classic model based on the Generalized Method of Moments (GMM), but it is necessary to use the Monte Carlo simulation method to perform the moment function in the application, so the degree of approximation has a greater impact on the effect of parameter calibration [9, 10]. Chen used the Method of Simulated Moments (MSM) to calibrate the parameters by minimizing the distance function between the moment of the simulated time series and the moment of the real data [11]. This process requires more calculations and different choices of moments will also have an impact on the results. Grazzini *et al.* introduced the Bayesian method to the ABM calibration to replace the simulated minimum distance method [12], but the work of Canova, Salaand, Fagiolo, and Roventini proved choosing a prior distribution may produce artificial curvature [13]. Francesco Lamperti *et al.* combined supervised machine learning and sampling methods to calibrate and verify the model. This also has certain requirements on the prior distribution [14]. Ye *et al.* investigated the state transition of the system from a macro perspective, and introduced the mean field approximation method to calibrate the micro parameters [15]. On this basis, this paper proposes to introduce Reinforcement Learning (RL) into the process of parameter calibration to replace the method of mean field approximation, and will take experiments to verify that the method can achieve high accuracy and low computational complexity.

This paper mainly provides the following three contributions:

The research was supported by National Key R&D Program of China (2018YFB1702700), National Natural Science Foundation of China (U1909204, 61773381, U1811463 & U19B2029), and Chinese Guangdong's S&T project (2019B1515120030, 2020B0909050001). Youth Innovation Promotion Association Chinese Academy of Sciences under Grant 2021130.Dongguan STS joint project 20201600200072.

1) Use RL to learn the behavior of the agent, and then simulate the behavior of agents, and use this to calibrate the state transfer parameters.

2) Propose to treat the interaction between agents as the parameters of the agent itself to reduce the computational complexity.

3) Through the application case study of population migration, it is proved that the method behave well in accuracy and computational complexity.

The rest of this paper is as follows: section II reviews some classical methods, and section III introduces method using RL, then section IV is about experiment, finally section VI is conclusion.

II. REVIEW OF METHODS OF CALIBRATION

A. Review of Classical Methods

Due to the complexity of the internal structure of multi-agent system (MAS), it is difficult to calibrate the parameters directly from the microscopic level. Therefore, in theory, it is feasible to minimize the error between the output of the ABM system and the statistical data of the actual system. This inspired the method based on moments. The idea of the Method of Simulated Moments (MSM) is to generate a distance function [11]. The parameters are optimized by minimizing the distance function in the whole parameter space. This process can be expressed as:

$$argmin\, Dist(X^R, X^S, \theta)$$

where, X^R stands for the sampling moment of real-world data, and X^S stands for the simulated moments. $Dist$ is the distance function of sampling moment and simulated moment. Search parameter θ to minimize the distance function in parameter space. But the choosing of moments has a greater impact on the results of this method and the method has high computational complexity.

Then Bayesian inference techniques was introduced to ABM calibration in Grazzini *et al.* [12]. The Bayes theorem is applied in calibration:

$$p(\theta|Y^R) \propto L(\theta; Y^R) \cdot p(\theta)$$

where, $L(\theta; Y^R) = p(Y^R|\theta)$ stands for the likelihood, $p(\theta)$ is the prior distribution of parameters and $p(\theta|Y^R)$ is the posterior distribution. Y^R is observed statistics. Bayesian method solves the parameter θ to maximize $p(\theta|Y^R)$.

The problem of moment-based methods that the calculation of distance functions need numerical approximation methods does not exist in Bayesian method any more. And the Bayesian method uses the information from the whole distribution but not from specific moments, which could make Bayesian method more efficient.

However, some steps involved in this frame need heavy computation. The complexity of agent based model needs efficient sampling techniques.

Francesco Lamperti *et al.* introduced machine learning and intelligent iterative sampling to ABM calibration [14]. This approach solves the problem of exploration in parameter space and calibration of ABM parameters through drawing pool of parameter combinations. So it can reduce computation time for parameter space exploration and calibration. However, the heuristic search is needed in decision tree surrogate so when the scale of parameters increasing, the computation of the method will become much more extensive.

Calibrating from a fundamental perspective in detail helps to get model results that have higher resolution and reference value. Yu et al. establish the AMETS (Agent-based Model for Emission Trading Scheme) model and calibrate two parts of system parameters [18]. The first part is calibrated with values of parameters from FORECAST model [19] and collected data. For calibration in the second part, they assume functions and calibrate based on data in detail. The detail data comes from FORECAST model and related works. This helps the results of AMETS achieve high accuracy.

B. Review of State Transfer

In Ye's work [15], the behavior of agents can be regarded as a high-order Markov process, so the change of the macroscopic state of the entire MAS can be expressed by a state transfer equation:

$$\begin{cases} x(t) = T(t) \cdot x(t-1) \\ x(0) = x_0 \end{cases} \quad (1)$$

where, $x(t) = [x_1(t), x_2(t) \cdots x_N(t)]^T$ stands for the system state, $x_i(t)$ stands for the number of agents with the i-th micro state at time t. N stands for the number of states. x_0 is the initial state of the system. $T(t)$ is the state transfer matrix. $T(t) \in R^{N\times N}$ and $T_{ij}(t)$ stands for the transfer probability from i-th micro state to j-th micro state at time t.

Let $y(t)$ stand for system observation. Considering the measurement is linear, then the M-dimensional vector $y(t)$ can be expressed as:

$$y(t) = B(t) \cdot x(t) \quad (2)$$

where, $B(t) \in R^{M\times N}$ stands for measurement matrix. Let $\hat{y}(t)$ stand for the actual observation from the realistic system, then the performance of the system can be expressed as:

$$J = \sum_{t=1}^{K} J(t) = \sum_{t=1}^{K} [y(t) - \hat{y}(t)]^T \cdot V \cdot [y(t) - \hat{y}(t)] \quad (3)$$

where, K stands for the number of steps, V is used to express the importance of each metric. Then $T(t)$ can be solved through minimizing J.

III. CALIBRATION USING REINFORCEMENT LEARNING

On the basis of macro state transfer, this paper introduces RL to calibration. The process is shown in Fig. 1. For all N agents that are with the i-th micro state at time $t-1$, system performs one step forward, and different agents will transfer to different states. Expressing the number of agents that transfer from i-th state to j-th state as n_{ij}, then the transfer probability from i-th state to j-th state at time t can be calculated with the equation:

$$T_{ij} = \frac{n_{ij}}{N} \quad (4)$$

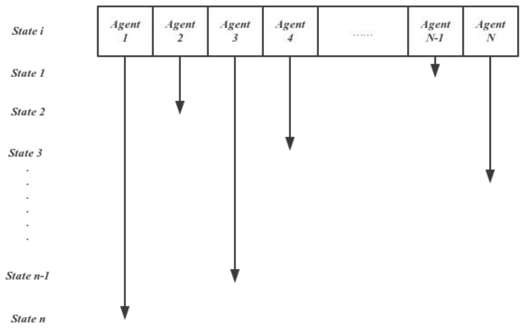

Fig. 1.State transfer after RL.

The step is decided by RL. After training, agents will choose the state that can obtain the most profit to transfer. Since the parameters of the agents are different, their transfer targets will be different.

For each agent, the factors that affect its state change are divided into: the agent's own parameters related to state changes, the influence from other agents, environmental factors (including random effects) and the agent's decision-making methods and parameters. Due to the heterogeneity of MAS, different agents will be affected by different other agents. If such impacts are considered in RL, a feasible method is to adopt Multi-agent RL, which can be more effective to simulate the complex network of relationships between agents. However, with the increasing of amounts of agents, the computational cost of this method has also become extremely high. Therefore, this paper expresses the influence from other agents in decision-making as the agent's own parameters, then the agent can be expressed as $a\{t, \theta_i(t), Nei(t), W_i, Env\}$, where $\theta_i(t)$ stands for agent's own parameters, $Nei(t)$ stands for the influence from other agents, W_i stands for the decision-making parameters and Env stands for the parameters from environment. So the state transfer probability T_{ij} can be used to calibrate parameters with the function:

$$T_{ij} = (a\{t, \theta_j(t), Nei(t), W_j, Env\} \\ |a\{t-1, \theta_i(t-1), Nei(t-1), W_i, Env\}) \quad (5)$$

The reward in RL is expressed as:

$$R = f(\theta_i(t), Nei(t), W_i(t), Env) \quad (6)$$

Through training, the decision-making process of the simulation system is constantly approaching the 'stylized facts'. The decision-making process of the agent is simulated at the micro level, then the state transfer probability is calibrated macroscopically after the decision is made. It is expected to improve the accuracy of ABM. The parameterized processing of the complex dynamic relationship network reduces the high computational complexity possibly.

IV. CASE STUDY: MIGRATION OF POPULATION

A. *Introduce of the Experiment*

This paper will verify the method proposed through the classic case of population migration. In order to facilitate comparison, we use the same data source as shown in [15, 17] for experiments. The experiment selects the annual population of China from 2000 to 2010 for simulation. There are three types of input data. The first type is the data of the fifth national census in 2000. The population in the data contains seven attributes, including gender, age, city of residence, ethnicity, registered provinces, marital status and birth status. The second type of input data is the classified population sample from the fifth national census. The sample contains 1,180,111 records, each of which covers personal and family attributes, and provides detailed information about a specific individual (private information is omitted). According to the final census data, the sample's data accounted for 0.95% of the national population. These two types of inputs are used to generate a basic comprehensive population in 2000. The third type of input data is the annual statistical data of 361 cities, including annual average income, birth rate, death rate and other demographic characteristics.

In the experiment, we set a Q table with a size of 361×361 (corresponding to the number of cities). The *i*-th row and *j*-column of the Q table represents the *q* value of the migration from city *i* to city *j*. Q table is used to train the decision-making process of the agent. After *n* iterations, the value of the Q table tends to converge. At this time, the city represented by the number of the column with the largest *q* value in the *i*-th row of the table will be the migration target city of the agent. The iteration equation of Q table is as follow:

$$Q(s, a) = (1-\sigma) \cdot Q(s, a) + \sigma \cdot (R + r \cdot maxQ(s')) \quad (7)$$

where, σ stands for the learning rate, $r \in (0,1)$ stands for the attenuation factor, $\max Q(s')$ stands for the future reward. R is the reward of action. The form of R can been expressed as:

$$R = k - (E_{or} - E_{des}) \quad (8)$$

where, k is an artificial constant. E_{or} means the evaluation of the value with origin state and E_{des} with destination state. The evaluation of the value is expressed as:

$$E = \alpha \cdot \theta_i(t) + \beta \cdot Nei(t) + \gamma \cdot W_i(t) \\ +(1-\alpha-\beta-\gamma) \cdot Env \quad (9)$$

where, α, β, γ stand for the weight of each parameter. The total of the 4 weights is 1.

B. *Experiment Results*

In the experiment, we set scale as 10,000, which means that one agent in simulation stands for 10,000 people in real statistics. The relative error is set to evaluate the accuracy of simulation. The relative error is expressed as:

$$error = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{n_{si} - n_{ai}}{n_{ai}}\right|$$

where N stands for the total number of cities, n_{si} is the number of population in the i-th city in simulation system. n_{ai} is the number of population in the i-th city in actual statistics. We compare the results of experiments with data shown in Ye's work [15]. The experiment results are shown in Fig. 2 and Table I.

As can be seen, the relative errors of our experiment range from 11% to 30%, and the errors of mean-field

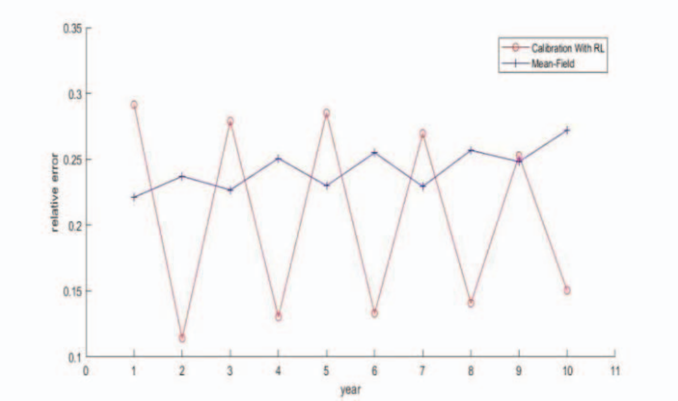

Fig. 2. Results of two methods of calibration

calibration range from 22 % to 28 %. Both methods get relative errors that are less than 30%, which means that two methods of calibration can do well in population migration experiment.

TABLE I. RELATIVE ERRORS OF TWO METHODS OF CALIBRATION

	2001	**2002**	**2003**	**2004**	**2005**
Calibration with RL	29.15%	11.40%	27.95%	13.03%	28.52%
Mean-field	22.11%	23.71%	22.67%	25.07%	22.99%
	2006	**2007**	**2008**	**2009**	**2010**
Calibration with RL	13.30%	26.96%	14.07%	25.28%	15.04%
Mean-field	25.50%	22.93%	25.68%	24.81%	27.20%

However, the relative errors of our method are much less than these of mean-field method in the year of 2002, 2004, 2006, 2008 and 2010. And the relative errors in the other years are close. This can be seen in Fig. 2 intuitively.

The running time of our methods in experiment is 2,257 seconds, and mean-field calibration needs 2,158 seconds. The difference in running time is only about 100 seconds. The reason is that we treat the interaction between agents as the parameters of the agent itself. This simplifies the complexity of the calculation to a certain extent.

V. CONCLUSION

This paper proposes to use RL in parameter calibration. RL is used to simulate the behavior of the agent to calibrate the state transfer probability, and then the microscopic parameters of the agent are calibrated according to the state transfer probability. When dealing with the complex interaction of agents, we propose to set a parameter to generalize the relationship. Then this paper proves our method can achieve high accuracy and similar computational complexity comparing to mean-field calibration through the experiment of population migration.

After getting state transfer probability, how to get the unique parameters is still a question. Now we can only limit a range of the concrete parameter, more methods to solve this problem are needed. Another problem is about the application of RL. RL is a powerful learning method, but how to better apply it to ABM calibration requires more work.

REFERENCES

[1] Jiaqi Wang, Andong Sheng, Dabo Xu, Zhiyong Chen and Youfeng Su, “Event-Based Practical Output Regulation for a Class of Multiagent Nonlinear Systems.” IEEE Transactions on Cybernetics, 2019, 49(10): pp. 3689-3698.

[2] Zhang, Huaipin, D. Yue, C. Dou, W. Zhao and Xiangpeng Xie, “Data-Driven Distributed Optimal Consensus Control for Unknown Multiagent Systems with Input-Delay.” IEEE Transactions on Cybernetics 49, 2019: pp. 2095-2105.

[3] P. Ye, X. Wang, C. Chen, Y. Lin and F. Wang, “Hybrid Agent Modeling in Population Simulation: Current Approaches and Future Directions.” Journal of Artificial Societies and Social Simulation, 2016: 19(1), 12.

[4] Chen, Zong-Gan, Zhi-hui Zhan, Ying Lin, Yue-jiao Gong, T. Gu, F. Zhao, Huaqiang Yuan, X. Chen, Q. Li and J. Zhang, “Multiobjective Cloud Workflow Scheduling: A Multiple Populations Ant Colony System Approach.” IEEE Transactions on Cybernetics 49, 2019: pp. 2912-2926.

[5] Quan Shi, Tieshan Li, Jingqi Li, C. L. Chen, Yang Xiao, Qihe Shan, “Adaptive Leader-following Formation Control with Collision Avoidance for a Class of Second-order Nonlinear Multi-agent Systems.” Neurocomputing, 2019, 350: pp. 282-290.

[6] Hu, H., Guanghui Wen, W. Yu, J. Cao and T. Huang, “Finite-Time Coordination Behavior of Multiple Euler − Lagrange Systems in Cooperation-Competition Networks.” IEEE Transactions on Cybernetics 49, 2019: pp. 2967-2979.

[7] Magliocca, N., V. McConnell and M. Walls, “Integrating Global Sensitivity Approaches to Deconstruct Spatial and Temporal Sensitivities of Complex Spatial Agent-Based Models.” J. Artif. Soc. Soc. Simul. 21, 2018.

[8] Fagiolo, G., M. Guerini, F. Lamperti, A. Moneta and A. Roventini, “Validation of Agent-Based Models in Economics and Finance.” 2017.

[9] Alfarano, S., T. Lux and F. Wagner, “Estimation of Agent-Based Models: The Case of an Asymmetric Herding Model.” Computational Economics 26, 2005: pp. 19-49.

[10] Fagiolo, G., M. Guerini, F. Lamperti, A. Moneta and A. Roventin, “Validation of Agent-Based Models in Economics and Finance.” 2017.

[11] Chen, S., Chia-Ling Chang and Ye-Rong Du, “Agent-based economic models and econometrics.” The Knowledge Engineering Review 27 2012: pp. 187-219.

[12] Grazzini, Jakob, Matteo G. Richiardi and M. Tsionas, “Bayesian Estimation of Agent-Based Models.” Journal of Economic Dynamics and Control 77, 2015: pp. 26-47.

[13] Canova, F. and L. Sala, “Back to Square One: Identification Issues in DSGE Models.” Macroeconomics: Aggregative Models eJournal 2005.

[14] Lamperti, F., A. Roventini and Amir Sani, “Agent-Based Model Calibration using Machine Learning Surrogates.” Sciences Po publications 2017.

[15] Peijun Ye,“Bridging the Micro and Macro: Calibration of Agent-Based Model Using Mean Field Dynamics.” Unpublished.

[16] Schelling, T., “Models of Segregation.” The American Economic Review 59, 1969: pp. 488-493.

[17] P. Ye, X. Hu, Y. Yuan and F.-Y. Wang, “Population Synthesis Based on Joint Distribution Inference without Disaggregate Samples.” Journal of Artificial Societies and Social Simulation, 2017, 20 (4): 16.

[18] Yu, S., Ying Fan, L. Zhu and W. Eichhammer, “Modeling the emission trading scheme from an agent-based perspective: System dynamics emerging from firms' coordination among abatement options.”Eur. J. Oper. Res. 286, 2020: pp. 1113-1128.

[19] Fleiter, T., M. Rehfeldt, A. Herbst, Rainer Elsland, Anna-Lena Klingler, P. Manz and S. Eidelloth, “A methodology for bottom-up modelling of energy transitions in the industry sector: The FORECAST model.” Energy Strategy Reviews 22, 2018: pp. 237-254.