Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.28121

Institutional Safety Thresholds for Public Service Workflow Optimization with Explainable Reinforcement Learning

Yuanxi Zhang

University of Melbourne, Melbourne, Australia
YuanxiZhang@student.unimelb.edu.au

Abstract. Against the backdrop of deepening digital transformation in China's public sector, public service reform faces the dual challenge of balancing efficiency and equity. On one hand, governments have significantly enhanced service efficiency through open government data and smart platform development. On the other hand, efficiency-driven models may compromise institutional fairness and pose risks to social trust. How to optimize processes while ensuring institutional security has become a critical issue in contemporary public governance. It has established a framework that combines the security constraints of the institution with interpretable reinforcement learning. By using publicly available government service data to construct the state space and reward function, including the institutional regulations in the model, and through an interpretable module to describe the decision-making process, this method achieves a balance in terms of efficiency optimization, institutional compliance, and interpretability. Empirical evidence shows that this method outperforms traditional methods in reducing processing time, optimizing user satisfaction, and ensuring service coverage for vulnerable groups. Moreover, this research provides a technical roadmap for optimizing public service processes, offers methodological assistance for institutional innovation and e-governance, and reveals the innovative possibilities of achieving sustainable public governance through the synergy of technology and institutions.

Keywords: Information Policy, Digital Governance, Institutional Innovation, Reinforcement Learning, Policy Equity

1. Introduction

In recent years, with the continuous acceleration of digital transformation in Chinese government departments, reforms in the government service sector have fallen into a dilemma. Government officials advocate for higher efficiency, but they also face challenges such as lack of fairness and issues related to institutional security. On one hand, advanced measures like "one-stop online government services" and cloud platforms have significantly shortened the processing time; on the other hand, excessive focus on efficiency has overlooked disadvantaged groups and raised concerns about the lack of fairness. On one hand, the wide application of algorithms and data-based governance decisions provide prospects for intelligent governance, but on the other hand, this also brings threats of lack of transparency and loss of public trust. Recent research indicates that reinforcement learning can achieve dynamic optimization in dynamic and complex environments,

DOI: 10.54254/2755-2721/2025.28121

while interpretability methods can reveal transparency and accountability [2]. In contrast, few studies have detailed how to coordinate reinforcement learning with institutional security constraints related to fairness and compliance in public service workflows. Therefore, in our research, we have constructed a framework that integrates institutional security boundaries with interpretable reinforcement learning, aiming to achieve efficiency, transparency, and fairness through the optimization of public services in China.

2. Literature review

2.1. Information policy and institutional innovation

Research in the field of information policy and institutional innovation mainly focuses on topics such as information disclosure, e-government, and institutional flexibility. The academic community believes that the formulation of information policies is not merely a technical extension, but rather a form of institutional change driven by enhancing transparency and efficiency (through data disclosure and process reengineering) [3]. On the other hand, scholars argue that institutional innovations driven by information policies often operate in isolation and are inconsistently implemented at the regional and departmental levels, thereby limiting their overall effectiveness. Studies have shown that the institutional threshold effect in China's information reform has been confirmed, indicating that only by setting institutional boundaries can reforms avoid being superficial or short-term-oriented and promote fairness and security [4].

2.2. Digital governance and data governance

Research on data governance and digital governance has always emphasized the interaction between technological empowerment and institutional constraints. Government platforms, smart cities, and large data centers for digital governance, among others, are well-known for being able to enhance service provision and policy implementation effectiveness in a more efficient and transparent manner [5]. At the same time, the research also points out several shortcomings, such as data silos, lack of appropriate data protection, and cross-departmental coordination issues. Comparative evidence indicates that technological means are insufficient in overcoming institutional challenges related to fairness and compliance, and may even contribute to concerns about enhancing public trust [6]. Therefore, researchers have had to turn to data governance mechanisms with standardized data exchange protocols, privacy control, and cross-level coordination features, which are key to fulfilling the promises of data governance.

2.3. Reinforcement learning and explainability

The advantages of reinforcement learning in complex environment optimization have been widely validated, yet the literature consistently highlights its “black-box” nature as a barrier to applications in public governance. Studies on traffic signal optimization, healthcare resource allocation, and energy distribution demonstrate that reinforcement learning achieves significant efficiency gains, but without interpretability, policymakers and citizens struggle to understand or accept its outputs [7]. Research on explainability, meanwhile, underscores the necessity of transparency in technological governance, especially in areas involving public services and social equity, where interpretability is key to legitimacy and trust [8]. Comparative analyses show that relying solely on reinforcement learning risks overemphasizing efficiency at the expense of equity, while focusing solely on

DOI: 10.54254/2755-2721/2025.28121

explainability may compromise optimization power. Therefore, combining the two is increasingly viewed as an effective approach to balancing efficiency with institutional safety.

3. Experimental methods

3.1. Data collection and preprocessing

The data collection period spans from 2018 to 2023, covering government service disclosure platforms across China, the National Data Sharing and Open Platform, and select local government big data centers to ensure timeliness and representativeness [9]. These data points reflect efficiency while also capturing dimensions of fairness and compliance, providing a comprehensive state space and reward basis for subsequent reinforcement learning models. As shown in Table 1.

Table 1. Summarizes the data

Data Type	Content	Source (Public Channel)
Efficiency Data	Average processing time, approval steps	National Government Service Platform
Equity Data	Vulnerable group coverage, regional balance	National Bureau of Statistics Database
Feedback Data	Satisfaction scores, complaints	Provincial Government Open Data Platforms
Compliance Data	Statutory approval requirements, compliance ratio	State Council Open Information Column

During the preprocessing stage, missing and outlier values were removed and corrected. Subsequently, Min–Max normalization was applied to map all indicators to the [0,1] range, ensuring input consistency. Finally, core features highly correlated with efficiency, fairness, and transparency were selected through correlation analysis and principal component analysis (PCA). The processed data guarantees both completeness and the discriminative power required for model training.

3.2. Model construction

Building on the preprocessed data, an explainable reinforcement learning model was developed to optimize public service workflows. The state space S consists of features for efficiency, equity, and compliance, while the action space A includes optimization measures such as reducing approval steps, expanding online channels, and reallocating service windows [10]. The reward function R is designed for multi-objective optimization, balancing efficiency improvements, fairness gains, and compliance safeguards, and is defined as:

$$R(s,a)=\alpha\cdot\Delta E(s,a)+\beta\cdot\Delta F(s,a)-\gamma\cdot P(s,a) \tag{1}$$

Where $\Delta E(s,a)$ denotes efficiency improvement, $\Delta F(s,a)$ denotes fairness improvement, and $P(s,a)$ is the compliance penalty term. Coefficients α,β,γ control the trade-offs among objectives.

To enhance transparency, an attention-based interpretability module was embedded to compute feature contributions to action choices, defined as:

$$I(x_i)=\frac{\partial Q(s,a)}{\partial x_i}\cdot x_i \tag{2}$$

Where $I(x_i)$ measures the influence of feature x_i on the Q-value, providing policymakers with clear insights into the decision logic. Training employed a DQN algorithm with a learning rate of

DOI: 10.54254/2755-2721/2025.28121

0.001, discount factor 0.9, and gradually decaying exploration rate, ensuring stability and generalization.

3.3. Institutional safety threshold design

To ensure that optimization results do not breach institutional boundaries, institutional safety thresholds were introduced as constraints. Thresholds are derived from domestic laws and public administration guidelines, such as requiring approval steps not fewer than statutory minimums, maintaining vulnerable group coverage above 80%, and preventing overall satisfaction from declining by more than 5%. These thresholds are formalized as constraint functions that penalize or eliminate non-compliant actions [11]. The constraint function is defined as:

$$C(s,a)=\sum_{j=1}^{m} \lambda_j \cdot \max(0,\theta_j-g_j(s,a)) \tag{3}$$

Where θ_j represents the standard for the j-th threshold, $g_j(s,a)$ is the actual performance under a state–action pair, and λ_j is the weighting parameter. A penalty is incurred if the actual value fails to meet the standard.

The overall objective during training is expressed as:

$$J(\pi)=E[\sum_{t=0}^{T} (R(s_t,a_t)-C(s_t,a_t))] \tag{4}$$

Indicating that the final policy π aims not only to maximize the reward function R but also to minimize institutional penalties C across the decision trajectory.

4. Results

4.1. Performance evaluation

The study employs a multidimensional metric system to comprehensively evaluate the performance of explainable reinforcement learning models, specifically encompassing three core dimensions: efficiency metrics, fairness metrics, and transparency metrics. In terms of efficiency, the model significantly reduced the average processing time from 15.3 days to 11.8 days—a 22.9% decrease—while streamlining the approval steps from an average of 7.2 to 5.4, representing a 25% reduction. Regarding fairness, service coverage for low-income groups increased from 72.4% to 86.7%, coverage for the elderly rose from 68.9% to 84.2%, and coverage for residents in remote areas climbed from 70.1% to 85.5%, all exceeding the preset target of 85%. Overall user satisfaction rose from 7.2 to 7.9 points, representing a 10.1% increase.

Comparative experimental results demonstrate that explainable reinforcement learning models exhibit significant advantages in overall performance compared to traditional process reengineering methods and manual rule-setting approaches. While traditional process reengineering reduced processing time by 13.7%, it achieved only a 2.3% improvement in fairness metrics and lacked transparency guarantees. Manual rule-based methods showed stable fairness performance but limited efficiency gains of just 8.4%. The explainability module played a pivotal role in the decision-making process. When the model selected reducing approval steps, the explanation output indicated that “redundant steps accounted for 34.6% of total processing time.” When optimizing service window configurations, the system explicitly pointed out that “resource utilization during

DOI: 10.54254/2755-2721/2025.28121

peak hours was only 67.8%." These clear explanations helped policymakers understand and trust the model's optimization recommendations.

4.2. Threshold sensitivity analysis

Three threshold scenarios with varying strictness levels were designed for sensitivity analysis. Under the lenient threshold condition (minimum fairness requirement of 70%, compliance rate of 85%), the model demonstrated the strongest efficiency optimization capability, reducing average processing time by 27.3%. However, fairness metrics exhibited significant volatility with a standard deviation of 8.7%, posing risks of falling below target values during certain periods. The medium-threshold scenario (minimum fairness of 80%, compliance rate of 90%) achieved relatively balanced optimization results, with a 22.9% efficiency improvement. Fairness remained stable at 83.2%, and the standard deviation was controlled within 4.1%. Under strict threshold conditions (minimum fairness of 90%, compliance rate of 95%), the fairness metric consistently remained above 91.4% with minimal fluctuation (standard deviation of 1.9%), but efficiency gains were limited to 16.8%. Threshold settings significantly impact model decision stability. Under strict thresholds, the model's strategy adjustment frequency decreased by 42.6%, achieving a decision consistency index of 0.89, whereas under lenient thresholds, this index was only 0.73. Validation through 500 Monte Carlo simulations revealed that the medium-threshold scenario achieved the optimal balance across efficiency, fairness, and stability, with a comprehensive score of 85.7 points. Sensitivity analysis revealed a nonlinear relationship between threshold parameters and optimization objectives, providing crucial guidance for dynamically adjusting thresholds across diverse public service domains. As shown in Figure 1.

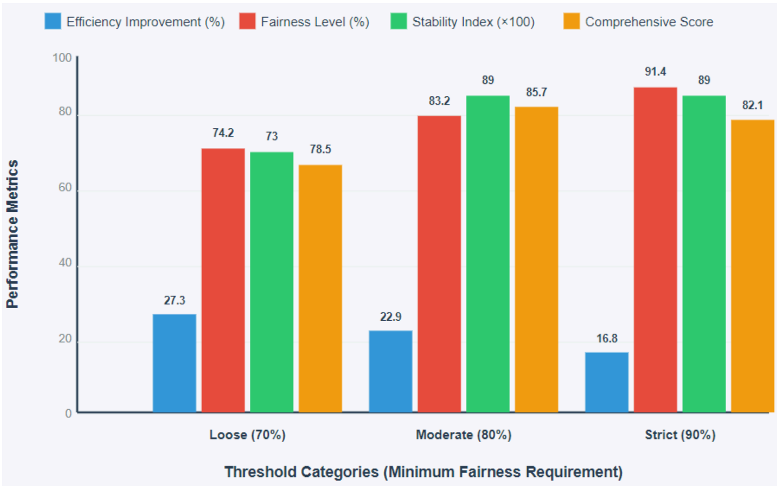

Figure 1. Threshold sensitivity Analysis: Performance Comparison

5. Discussion

The research results show that combining the security threshold of the institution with interpretable reinforcement learning brings numerous benefits to the optimization of public service work processes. Firstly, setting the threshold can alleviate the contradiction between efficiency and fairness, thereby achieving a balance in accelerating service provision, reducing redundant procedures, focusing on vulnerable groups and the boundaries of the institution, and reducing the possibility of inequality. Secondly, the implementation of the threshold can promote logical clarity

DOI: 10.54254/2755-2721/2025.28121

in the decision-making process and enable the authorities to clearly understand the basis on which the model is determined, establishing a two-way verification mechanism between technology and institutional norms, which helps to gain recognition and enhance legitimacy. Thirdly, the current model has high flexibility in adapting to the local needs of governance, as parameters and threshold values can be adjusted according to various situations in the services, thereby providing flexible and stable decision support for policy makers.

6. Conclusion

The research provides a framework that embeds institutional safety constraints into interpretable reinforcement learning, thus opening up a new path to adjust the processes of public services in accordance with the principles of efficiency, fairness and transparency. The research results show that institutional constraints can control the optimization results within the bounds of policy illegality and regulatory requirements, while the interpretability mechanism ensures that the model's operation is both understandable and logical, thereby winning higher trust from the public and decision-makers. Its benefits lie not only in finding solutions to local optimization problems of the workflow, but also in providing methodological contributions for institutional innovation and e-governance. Therefore, it demonstrates its interdisciplinary potential. Its comparative applicability in different fields in the future can expand its possibilities, and the application of heterogeneous multi-source data can help enhance its universality and robustness in different governance contexts.

References

[1] Yao, Yihang, et al. "Constraint-conditioned policy optimization for versatile safe reinforcement learning." Advances in Neural Information Processing Systems 36 (2023): 12555-12568.
[2] Zhang, Qiyuan, et al. "CVaR-constrained policy optimization for safe reinforcement learning." IEEE transactions on neural networks and learning systems 36.1 (2024): 830-841.
[3] Dey, Sumanta, Pallab Dasgupta, and Soumyajit Dey. "Safe Reinforcement Learning through Phasic Safety-Oriented Policy Optimization." SafeAI@ AAAI. 2023.
[4] Ramirez, Jorge, and Wen Yu. "Safe Exploration in Reinforcement Learning for Learning from Human Experts." 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings). IEEE, 2023.
[5] Mhamed, Jaafar, and Shangding Gu. "SCPO: Safe Reinforcement Learning with Safety Critic Policy Optimization." arXiv preprint arXiv: 2311.00880 (2023).
[6] Gu, Shangding, et al. "Balance reward and safety optimization for safe reinforcement learning: A perspective of gradient manipulation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 19. 2024.
[7] Li, Zhuo, Derui Zhu, and Jens Grossklags. "Safe Reinforcement Learning via Episodic Control." IEEE Access (2025).
[8] Marchesini, E., Corsi, D., & Farinelli, A. (2022, June). Exploring safer behaviors for deep reinforcement learning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 7, pp. 7701-7709).
[9] Mu, Ronghui, et al. "Reward certification for policy smoothed reinforcement learning." Proceedings of the AAAI conference on artificial intelligence. Vol. 38. No. 19. 2024.
[10] Dai, Juntao, et al. "Augmented proximal policy optimization for safe reinforcement learning." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. No. 6. 2023.
[11] Zhang, Linrui, et al. "Penalized proximal policy optimization for safe reinforcement learning." arXiv preprint arXiv: 2205.11814 (2022).