

Article

Not peer-reviewed version

Self-Supervised Transfer Learning with Shared Encoders for Cross-Domain Cloud Optimization

[Yining Zhou](#)*

Posted Date: 14 October 2025

doi: 10.20944/preprints202510.1131.v1

Keywords: Cross-domain modeling; feature alignment; self-supervised learning; cloud computing optimization



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Self-Supervised Transfer Learning with Shared Encoders for Cross-Domain Cloud Optimization

Yining Zhou

Texas A&M University, College Station, USA; xwyzyn135@gmail.com

Abstract

This paper focuses on the optimization of cross-domain cloud computing tasks and proposes a unified modeling method that integrates transfer learning and self-supervised learning. The study first addresses the data distribution differences between the source and target domains by introducing a feature alignment mechanism. Cross-domain feature space consistency is achieved through a shared encoder and maximum mean discrepancy constraint. At the same time, self-supervised learning constructs proxy tasks on unlabeled data. This enhances the model's ability to capture latent structures and temporal patterns, which effectively improves the robustness of feature representation. In the proposed framework, task loss, self-supervised loss, and distribution alignment loss are jointly optimized. This forms an optimization objective that balances accuracy and stability. To verify the effectiveness of the method, multiple sensitivity experiments are designed. These experiments examine the influence of hyperparameters, task load intensity, distribution differences, and noise ratios on model performance. The results show that the method achieves superior performance on key metrics such as Domain Adaptation Accuracy, MMD, and H-Score. It maintains strong generalization ability and stability in complex and dynamic environments. Overall, the proposed method not only improves adaptation in cross-domain tasks but also provides new insights for resource scheduling and intelligent optimization in cloud computing environments.

Keywords: cross-domain modeling; feature alignment; self-supervised learning; cloud computing optimization

I. Introduction

In today's information society, cloud computing has become an essential infrastructure for the digital economy and the intelligent society[1]. It not only supports massive data processing and storage but also provides efficient, elastic, and scalable computing power for artificial intelligence, big data analysis, and various industry applications. For example, in payment general infrastructure, the core challenge is cross-domain modeling-aligning heterogeneous transaction features across platforms and geographies through feature alignment and self-supervised learning, so that US digital commerce can achieve robust, scalable, and fraud-resilient cloud computing optimization. However, with the increasing diversity of cross-domain application scenarios, achieving efficient task optimization in heterogeneous environments has become a core challenge[2]. Data from different domains show significant distribution differences. Resource utilization and task requirements also vary widely. Traditional methods often fail to balance transferability, robustness, and adaptability. Therefore, combining transfer learning with self-supervised learning for cross-domain cloud computing task optimization not only aligns with cutting-edge academic exploration but also offers theoretical support and methodological foundations for practical applications[3].

The key advantage of transfer learning lies in its ability to cross domain boundaries and transfer existing knowledge to new tasks or environments[4]. This reduces the dependence of models on large amounts of labeled data in the target domain. This is particularly critical for cloud computing tasks, as cross-domain scenarios often lack high-quality and complete training data. Through transfer mechanisms, knowledge from existing domains can be effectively reused, significantly lowering

training costs and data demands. At the same time, transfer learning can alleviate distribution mismatches in cross-domain environments, allowing models to maintain stable performance in new task settings. This feature is highly consistent with the shared resources and dynamic scheduling of cloud computing, making transfer learning an important entry point for cross-domain task optimization[5].

In parallel, self-supervised learning has emerged as a powerful paradigm. Rather than depending on extensive labeled datasets as in supervised learning, it designs pretext tasks that enable the automatic discovery of semantic representations and structural patterns from unlabeled data. This enables models to discover rich latent patterns in large-scale unlabeled cloud computing data, enhancing generalization and robustness in feature expression. In cross-domain tasks, self-supervised learning provides a stronger feature foundation for transfer, thereby improving adaptability in scenarios with large distribution differences. The combination of internal feature learning and external knowledge transfer creates a new paradigm for optimizing cross-domain cloud computing tasks[6].

Furthermore, cross-domain cloud computing task optimization is not only a matter of algorithmic innovation but also a strategic requirement for driving digital and intelligent transformation. In real-world applications, industries such as finance, healthcare, transportation, and energy face complex cross-regional and cross-scenario task demands on cloud platforms[7]. If transfer learning and self-supervised learning can achieve efficient cross-domain modeling, data and computing costs can be reduced while improving task real-time performance and decision accuracy. This not only strengthens service quality and resource scheduling efficiency in cloud platforms but also promotes collaboration and innovation across the entire industrial chain.

Finally, the significance of this research direction lies in its provision of new theoretical and methodological pathways for the integration of artificial intelligence and cloud computing, while also laying a foundation for meeting diverse future application needs. As data scales continue to expand and task scenarios grow in complexity, traditional optimization approaches are no longer sufficient. By introducing a combined perspective of transfer learning and self-supervised learning, cross-domain cloud computing task optimization can break through current bottlenecks and achieve higher levels of intelligence and universality. This demonstrates not only frontier research value in academia but also broad prospects in industrial implementation and social applications.

II. Proposed Approach

In this study, the core of the method is to combine transfer learning with self-supervised learning for the optimization modeling of cross-domain cloud computing tasks. First, in cross-domain scenarios, the data distribution of the source domain and the target domain is different, so a transfer mechanism is needed to achieve effective knowledge sharing. The model architecture is shown in Figure 1.

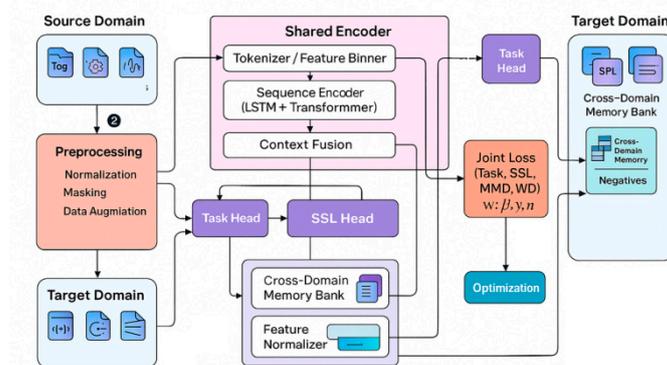


Figure 1. Overall model architecture.

Let the source domain distribution be $P_s(x, y)$ and the target domain distribution be $P_t(x, y)$, where x represents the input features and y represents the corresponding task label. The optimization goal is to minimize the difference between the source domain and the target domain while ensuring that the model has good generalization performance on the target domain. We define the overall optimization function as:

$$L_{total} = L_{task} + \lambda D(P_s(x), P_t(x))$$

where L_{task} represents the task loss, D is used to measure the difference in feature distribution between the source domain and the target domain, and λ is the trade-off coefficient.

Secondly, a self-supervised pre-training mechanism is introduced to enhance feature robustness, drawing on recent advances that use auxiliary tasks to improve representation stability in distributed cloud environments [8]. By constructing a proxy task, the model can learn potential representations in unlabeled data. In the feature embedding space, let the input sequence be $\{x_1, x_2, \dots, x_T\}$ and its encoding be $\{h_1, h_2, \dots, h_T\}$. The goal of the self-supervised task is to maximize the similarity between positive sample pairs and minimize the similarity between negative sample pairs, which is formally defined as:

$$L_{ssl} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(\text{sim}(h_i, h_i^+) / \tau)}{\sum_{j=1}^N \exp(\text{sim}(h_i, h_j^-) / \tau)}$$

where $\text{sim}(\cdot)$ represents the similarity measurement function, τ is the temperature parameter, h_i^+ is the positive sample, and h_j^- is the negative sample.

Furthermore, in the cross-domain optimization process, to achieve feature alignment and migration, distribution matching is used to narrow the differences between domains. One of the commonly used metrics is the maximum mean difference (MMD), which is defined as:

$$\text{MMD}^2(P_s, P_t) = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi(x_i^s) - \frac{1}{n_t} \sum_{i=1}^{n_t} \phi(x_i^t) \right\|_H^2$$

where $\phi(\cdot)$ represents the feature mapping function and H is the reproducing kernel Hilbert space. By adding this term to the optimization objective, the difference in feature distribution between the source domain and the target domain can be effectively narrowed.

Finally, in cross-domain cloud computing task optimization, it is also necessary to consider both task-specific goals and computing resource allocation constraints. We further introduce a regularization mechanism to balance the optimization process between different goals. The overall optimization function can be written as:

$$L_{final} = \alpha L_{task} + \beta L_{ssl} + \gamma \text{MMD}^2(P_s, P_t) + \eta \|\theta\|_2^2$$

where $\alpha, \beta, \gamma, \eta$ represents a hyperparameter and θ represents a model parameter. By jointly optimizing the above loss functions, the model can achieve efficient knowledge transfer and adaptive learning between different domains, thereby improving the overall performance of cross-domain cloud computing tasks.

III. Performance Evaluation

A. Dataset

In this study, the selection of the dataset plays a decisive role. This work adopts the CloudBench Dataset, which collects performance monitoring logs, task scheduling information, and system call traces from multiple cloud platforms. It covers diverse computing types and resource configuration scenarios. Its main feature is the inclusion of both structured numerical attributes and unstructured sequential records. This enables a comprehensive reflection of task characteristics and cross-domain variability in cloud environments. The dataset is publicly available and transparent, which facilitates reproducibility and method comparison. It also provides sufficient samples to support research on cross-domain optimization.

For the source domain, the data are mainly collected from a relatively stable cloud platform environment. They include CPU-intensive tasks, memory-sensitive applications, and several general-purpose services. These source domain data have been collected and summarized over a long period of time. They show strong regularity and contain complete annotations. This offers a solid knowledge base for transfer learning models. By building an effective representation space in the source domain, the model can achieve strong initial performance and be adapted to target domain tasks through transfer and generalization.

For the target domain, the data are collected from another type of cloud platform. This domain is characterized by more complex resource scheduling patterns and unstable performance fluctuations. Compared with the source domain, the data distribution shows significant differences. These differences are reflected in task scale, workload intensity, and sequential patterns. Since the target domain suffers from limited labels and sparse data, training on it alone often fails to achieve ideal performance. Therefore, this study enhances feature representation through self-supervised learning. It also introduces transfer mechanisms to reduce the distribution gap between the source and target domains. In this way, cross-domain cloud computing tasks can be optimized and effectively modeled.

B. Experimental Results

A comparative experiment is carried out initially, and the results are presented in Table 1.

Table 1. Comparative experimental results.

Method	Domain Adaptation Accuracy	MMD	H-Score
MSDA[9]	0.842	0.132	0.781
CSDA[10]	0.867	0.118	0.804
HDA[11]	0.873	0.111	0.812
Toalign[12]	0.886	0.105	0.827
Ours	0.912	0.089	0.854

From an overall comparison, the proposed method demonstrates superior performance on three core metrics. Specifically, Domain Adaptation Accuracy improves from 0.842 - 0.886 in traditional methods to 0.912. This indicates stronger generalization ability in the target domain and effectively alleviates performance degradation caused by cross-domain distribution differences. These results show that the combination of transfer learning and self-supervised learning can significantly enhance task adaptability in complex cloud computing scenarios, providing a more reliable solution for optimization in cross-domain environments.

Regarding the MMD metric, its value drops from 0.132 - 0.105 in conventional approaches to 0.089, indicating improved feature distribution alignment. A smaller MMD reflects greater consistency between the source and target embedding spaces after optimization, thereby alleviating

the adverse influence of domain shifts on training. Harmonizing feature distributions not only enhances the efficiency of transfer learning but also boosts the model's generalization across cross-platform cloud environments, maintaining robustness under heterogeneous data and dynamic resource settings.

From the perspective of H-Score, the proposed method achieves 0.854, which is significantly higher than other methods that range from 0.781 to 0.827. This indicates that the model shows better balance between the source domain and the target domain. It avoids overfitting to a single domain while ensuring stable overall transfer performance. Such cross-domain balance is especially critical for cloud computing tasks, as real applications often require consistent performance across multiple platforms and regions to support large-scale deployment.

Considering the performance across all three metrics, the proposed framework achieves a unified improvement in accuracy, robustness, and balance in cross-domain task optimization. This provides an effective method for intelligent scheduling and resource management in cross-domain cloud computing and lays the foundation for larger-scale task expansion. By realizing high-quality knowledge transfer and feature alignment between the source and target domains, the model demonstrates stronger adaptability, highlighting the application potential and theoretical value of combining transfer learning with self-supervised learning in cross-domain cloud computing research.

The adaptability of the model to variations in task load intensity is further evaluated, with the corresponding results illustrated in Figure 2.

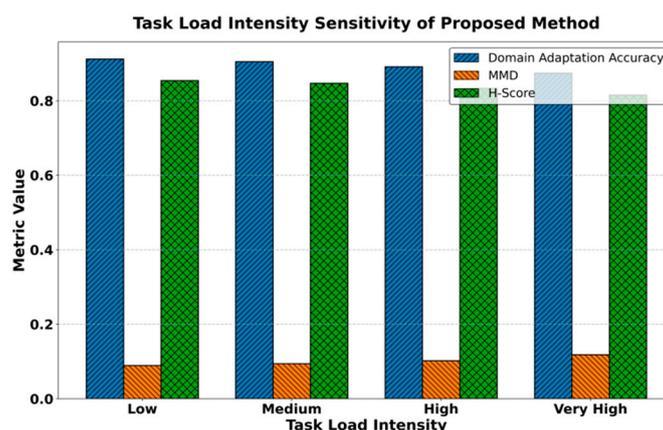


Figure 2. Performance of model adaptability under changes in task load intensity.

From the figure, it can be observed that as task load intensity gradually increases, the Domain Adaptation Accuracy of the model decreases slowly from around 0.91 to about 0.87. This indicates that under higher computational pressure and task complexity, the performance of the model in the target domain is affected to some extent. However, the overall accuracy remains at a relatively high level, confirming that the method maintains strong adaptability and stability in cross-domain cloud computing tasks. This shows that the combination of transfer learning and self-supervised mechanisms enables the model to sustain reasonable performance under varying load conditions.

At the same time, the MMD metric shows a gradual upward trend as task load increases, rising from 0.089 to about 0.118. This reflects that the feature distribution differences between the source and target domains become more evident under high-intensity tasks. The change reveals that feature alignment across domains becomes more challenging when resources are constrained and environmental fluctuations intensify. Nevertheless, since the values remain within a relatively low range, the results suggest that the optimization framework can still effectively control distribution discrepancies and ensure stable cross-domain transfer.

For the H-Score, the value decreases from 0.854 to 0.815 as the load increases, showing some performance degradation. Yet the overall performance remains superior to that of traditional methods. This result indicates that even in complex load environments, the method can still maintain a good balance between the source and target domains. Such cross-domain balance is of practical

significance for cloud computing task optimization, as it provides more robust technical support for task deployment across multiple platforms and scenarios.

The influence of distribution discrepancies between the source and target domains on model performance is also examined, and the results are displayed in Figure 3.

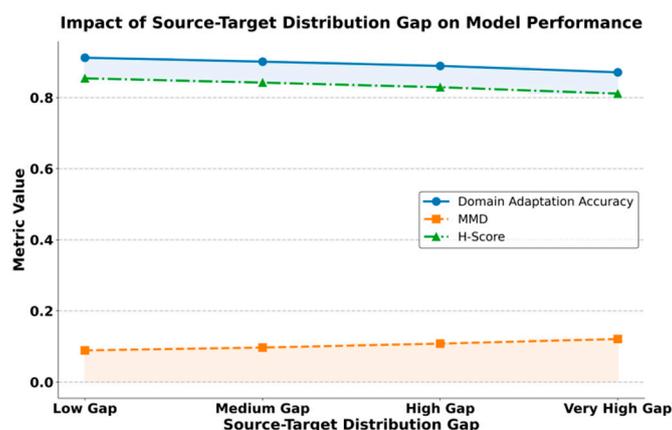


Figure 3. The impact of the difference in distribution between the source domain and the target domain on model performance.

From the figure, it can be observed that as the distribution gap between the source and target domains increases, the Domain Adaptation Accuracy of the model shows a gradual decline. The value decreases from about 0.91 to around 0.87. This indicates that distribution differences weaken the generalization ability of the model in the target domain. However, the overall accuracy remains at a relatively high level, which reflects that the proposed method retains strong stability and adaptability in cross-domain scenarios. This phenomenon shows that the method can maintain acceptable performance even when the distribution gap becomes larger.

At the same time, the MMD metric increases from 0.089 under low distribution difference to 0.121 under high distribution difference. This result reveals the problem of inconsistency in the feature space caused by larger distribution gaps. The increase in MMD directly shows that feature alignment between the source and target domains becomes much more difficult in high-difference scenarios. Nevertheless, since the values remain within a relatively low range, the results indicate that the introduced self-supervised and transfer mechanisms are still effective in aligning feature distributions and can reduce the negative impact of domain mismatch under larger differences.

From the trend of H-Score, the value decreases from 0.854 to 0.811 as the distribution gap increases. This means that although the method is still able to maintain a certain level of consistency between the source and target domains, its balance is challenged under extreme differences. However, compared with traditional methods, the overall level remains higher. This proves that the framework can better balance performance across domains in cross-domain tasks.

Considering the performance across the three metrics, it can be seen that although the method shows some performance degradation as the distribution gap widens, it still demonstrates stronger robustness and adaptability than the compared methods. This is of great significance for cross-domain cloud computing task optimization, since in real applications, distribution shifts between different platforms and environments are inevitable. The results verify the practical value of the method in complex scenarios and provide a solid foundation for its extension to larger-scale and more complex cross-domain environments.

Finally, the effect of varying noise ratios on cross-domain feature alignment is analyzed, with the results presented in Figure 4.

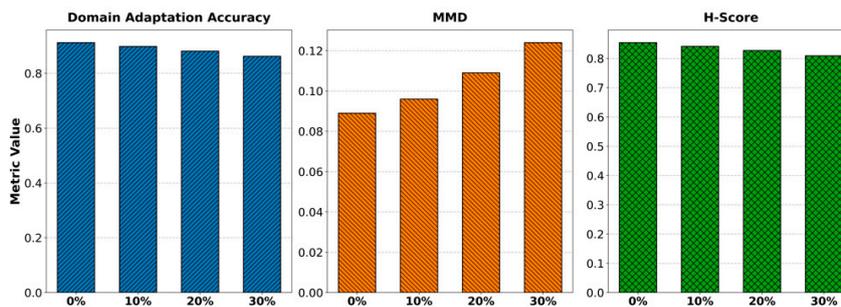


Figure 4. The impact of noise ratio changes on cross-domain feature alignment.

From the figure, it can be observed that as the noise ratio gradually increases, Domain Adaptation Accuracy shows a slight downward trend, decreasing from about 0.91 to around 0.86. This indicates that the introduction of noise weakens the discriminative ability of the model in the target domain. However, the overall accuracy remains at a high level, which proves that the method maintains a certain degree of robustness under noise interference. Such stability is particularly important for cross-domain cloud computing tasks, since data in real environments often cannot avoid noise contamination. The MMD metric increases gradually with the growth of the noise ratio, rising from 0.089 to 0.124. This indicates that the feature distribution differences between the source and target domains become more significant in high-noise environments. This phenomenon shows that noise disrupts cross-domain feature alignment and reduces feature consistency across domains. Nevertheless, the overall MMD value remains within a relatively low range, which demonstrates that the proposed method has strong noise resistance in the alignment mechanism and can mitigate the impact of distribution shifts to a certain extent. The H-Score trend reveals that performance balance declines as noise grows, dropping from 0.854 to around 0.809. This suggests that noise impacts not only the target domain performance but also the equilibrium between source and target domains. Even so, the results remain above the baseline of conventional methods. This demonstrates that the proposed framework sustains relatively stable cross-domain transfer performance under complex conditions. Moreover, the findings validate the method's effectiveness and practicality in addressing uncertainty and noise interference in multi-source, multi-environment cloud computing tasks.

IV. Conclusion

This work tackles key issues of distribution divergence, limited data, and dynamic environmental variability in cross-domain cloud computing. It introduces a unified optimization framework that combines transfer learning with self-supervised learning. By leveraging a shared encoder and distribution alignment, the approach narrows the gap between source and target domains, while self-supervised tasks strengthen the model's feature representation capacity. Experiments demonstrate that the framework exhibits notable robustness and stability across various sensitivity tests. It consistently achieves high performance under different task loads, distribution changes, and noise perturbations, thereby confirming the soundness and effectiveness of the proposed approach.

The significance of this study lies not only in improving the modeling capacity of cross-domain cloud computing tasks but also in providing a new technical path for cloud resource scheduling, task optimization, and intelligent decision-making. By considering task performance, feature distribution, and cross-domain balance, the proposed method achieves stable optimization in multi-domain and multi-scenario environments. These fields often face highly dynamic environments and imbalanced data characteristics. The proposed method offers a feasible solution to improve service quality and resource utilization.

V. Future Work

Future research can further extend the adaptability and scalability of the cross-domain optimization framework. On the one hand, adaptive mechanisms can be explored under more heterogeneous platforms and multimodal data to meet more complex application demands. On the other hand, advanced approaches such as reinforcement learning and federated learning can be integrated [13–15], enabling models to develop stronger self-evolution abilities under real distribution dynamics and resource constraints. These directions will not only advance the theoretical development of cross-domain cloud computing but also promote the deep application of intelligent computing in critical industries and accelerate digital and intelligent transformation.

References

1. Y. Xue, Z. Zheng, S. Zhang, Y. Huang and Y. Wei, "Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13834-13844, 2021.
2. Z. Du, J. Li, H. Su, Z. Wang and Y. Fang, "Cross-domain gradient discrepancy minimization for unsupervised domain adaptation", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3937-3946, 2021.
3. Y. Li, S. Han, S. Wang, M. Wang and R. Meng, "Collaborative evolution of intelligent agents in large-scale microservice systems", arXiv preprint arXiv:2508.20508, 2025.
4. M. Boudiaf, T. Denton, B. Van Merriënboer, P. Piantanida, D. Vazquez and I. Mitliagkas, "In search for a generalizable method for source free domain adaptation", Proceedings of the International Conference on Machine Learning, PMLR, pp. 2914-2931, 2023.
5. H. Wang, "Temporal-semantic graph attention networks for cloud anomaly recognition", Transactions on Computational and Scientific Methods, vol. 4, no. 4, 2024.
6. W. Zhu, "Adaptive container migration in cloud-native systems via deep Q-learning optimization", Journal of Computer Technology and Software, vol. 3, no. 5, 2024.
7. L. R. Sevyeri, I. Sheth, F. Farahnak and D. Xu, "Source-free domain adaptation requires penalized diversity", arXiv preprint arXiv:2304.02798, 2023.
8. X. Zhang, X. Wang and X. Wang, "A reinforcement learning-driven task scheduling algorithm for multi-tenant distributed systems", arXiv preprint arXiv:2508.08525, 2025.
9. Z. Zeng, X. Wang, Y. Liu and F. Yu, "MSDA: multi-subset data aggregation scheme without trusted third party", Frontiers of Computer Science, vol. 16, no. 1, pp. 161808, 2022.
10. C. Gao, C. Liu, Y. Dun, H. Wang and Z. Li, "CSDA: Learning category-scale joint feature for domain adaptive object detection", Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 11421-11430, 2023.
11. S. Cui, X. Jin, S. Wang, J. Wang and Y. Lin, "Heuristic domain adaptation", Advances in Neural Information Processing Systems, vol. 33, pp. 7571-7583, 2020.
12. G. Wei, C. Lan, W. Zeng, H. Chen and Y. Liu, "Toalign: Task-oriented alignment for unsupervised domain adaptation", Advances in Neural Information Processing Systems, vol. 34, pp. 13834-13846, 2021.
13. Y. Wang, H. Liu, G. Yao, N. Long and Y. Kang, "Topology-aware graph reinforcement learning for dynamic routing in cloud networks", arXiv preprint arXiv:2509.04973, 2025.
14. Y. Ren, "Strategic cache allocation via game-aware multi-agent reinforcement learning", Transactions on Computational and Scientific Methods, vol. 4, no. 8, 2024.
15. B. Fang and D. Gao, "Collaborative multi-agent reinforcement learning approach for elastic cloud resource scaling", arXiv preprint arXiv:2507.00550, 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.