

Исследование эффективности масштабирования алгоритма многоагентного обучения с подкреплением для балансировки трафика

Озерова Дарья Александровна¹, Степанов Евгений Павлович²

¹ Кафедра автоматизации систем вычислительных комплексов, e-mail: s02190164@gse.cs.msu.ru

² Кафедра автоматизации систем вычислительных комплексов, e-mail: estepanov@lvk.cs.msu.ru

Инженерия трафика (Traffic Engineering, TE) — хорошо зарекомендовавший себя механизм, который играет фундаментальную роль в производительности современного Интернета. Его основной целью является обеспечение эффективной и надежной работы сети при условии оптимизации использования сетевых ресурсов. Одной из реализаций TE является алгоритм балансировки трафика, критерием оптимизации которого является минимизация максимальной нагрузки на сетевой канал. В данном алгоритме реализована архитектура, сочетающая в себе многоагентное обучение с подкреплением (Multi-Agent Reinforcement Learning, MARL), графовые нейронные сети (Graph neural networks, GNN) и консистентное хеширование (Consistent Hashing, CH). На начальном этапе алгоритма запускается графовая нейросеть MPNN (Message Passing Neural Network), которая принимает на вход сетевую топологию в виде ориентированного мультиграфа и возвращает его векторное представление, где на выходе для каждого узла будет сформирован собственный вектор значений, который в дальнейшем будет использоваться для формирования хеш-весов для ребер сетевой топологии.

Недостаток алгоритма заключается в сложности масштабирования данной модели с точки зрения роста размера топологии сети. Работа MPNN на больших сетевых топологиях требует большого объема оперативной памяти для хранения векторного представления полного графа на каждой итерации и характеризуется долгим временем выполнения.

Данная работа посвящена разработке и реализации модуля динамического масштабирования вышеописанного алгоритма. Для минимизации использования ресурсов вычислительной системы было принято решение распараллелить работу MARL-GNN-CH алгоритма на разных частях сетевой топологии. Для решения поставленной задачи был проведен обзор алгоритмов разбиения графа на подграфы, основными критериями которого были сбалансированность разбиения и минимизация размера среза. Также было проведено функциональное и нагрузочное тестирование оптимизированного алгоритма и откалиброваны значения метрик масштабирования.

СПИСОК ЛИТЕРАТУРЫ

- [1] Bernárdez, G., Suárez-Varela, J., López, A., Wu, B., Xiao, S., Cheng, X., ... Cabellos-Aparicio, A. (2021, November). Is machine learning ready for traffic engineering optimization?. In 2021 IEEE 29th International Conference on Network Protocols (ICNP) (pp. 1-11). IEEE.
- [2] Kernighan, B. W., Lin, S. (1970). An efficient heuristic procedure for partitioning graphs. The Bell system technical journal, 49(2), 291-307.
- [3] Gupta, J. K., Egorov, M., Kochenderfer, M. (2017). Cooperative multi-agent control using deep reinforcement learning. In Autonomous Agents and Multiagent Systems: AAMAS 2017 Workshops, Best Papers, São Paulo, Brazil, May 8-12, 2017, Revised Selected Papers 16 (pp. 66-83). Springer International Publishing.
- [4] Zeng, H., Zhou, H., Srivastava, A., Kannan, R., Prasanna, V. (2019). Graphsaint: Graph sampling based inductive learning method. arXiv preprint arXiv:1907.04931.
- [5] Chiang, W. L., Liu, X., Si, S., Li, Y., Bengio, S., Hsieh, C. J. (2019, July). Cluster-gcn: An efficient algorithm for training deep and large graph convolutional networks. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery data mining (pp. 257-266).