Supercomputing for the Multi-Driver Routing

Extended Abstract

Zeyang Ye
Department of Applied Mathematics and Statistics
Stony Brook, NY
zeyang.ye@stonybrook.edu

Yuefan Deng (Advisor)
Department of Applied Mathematics and Statistics
Stony Brook, NY
yuefan.deng@stonybrook.edu

CCS CONCEPTS

• **Computing methodologies → Parallel computing methodologies**; • **Information systems → Mobile information processing systems**; *Data mining*;

KEYWORDS

Supercomputing, data mining, route recommendation, global optimization

1 PROBLEM STATEMENT

We extract 21,824 pick-up points from 12,481 cab mobility traces in Beijing, China. Then we calculate the probability of successful pickup of a passenger by a taxi at each pickup point. With this information, the potential traveling distance (PTD) for a route can be determined. The two criteria, distance and probability, are combined into one PTD value [1]. The new multi-driver routing problem we develop considers a set of routes. It aims to solve multi-driver PTD (MPTD) function for one additional requirement, that the routes should not overlap.

2 SIGNIFICANCE AND CHALLENGES

The rise of on-demand mobility systems becomes more and more significant. If the current trend continues, the taxis services, including those offered by Uber and Lyft, could be the major component of the traffic. On the other hand, traffic is now a major source of global greenhouse gas emissions [4]. A recent study [4] shows that a near optimal taxi service in New York City, if achieved, will lead to 30 percent reduction for the fleet size of the taxis. The multi-driver routing problem is designed to reduce the cruising time for the taxi drivers and contribute to achieving the optimal taxi services, which will bring more profit to the driver, make the passengers more convenient, reduce the traffic, and benefit the environment.

It is challenging to solve this problem. First, the data is large. Around the end of 2016, in New York City streets, including yellow cabs, Uber, and Lyft, there are over 60,000 taxi drivers and 600,000 daily trips [3]. Solving the multi-driver routing problems for high dimensions needs extremely long time. Second, the problem should be solved in a short time frame. The taxi driver needs the recommendation immediately and the traffic changes rapidly. The recommended route based on the current traffic can be quickly out-of-date.

3 NOVEL METHODOLOGIES

Simulated annealing (SA) is a stochastic optimization method for global optimization of non-convex objective functions. In each iteration, it proposes a move by perturbing its current solution and accepts the move with a probability. For parallel SA (PSA), each computing core performs a local SA for a few steps. When the number of independent steps reaches the prescribed mixing period, the local SAs mix their results by a prescribed mixing pattern that is measured by surviving rate, the ratio between the numbers of the surviving states and all states.

For PSA with domain decomposition (PSAD), each computing core runs SA global search to optimize one or a small number of routes with a subset of the pickup points. To make each route have the access to all the pickup points, after a preset number of steps, all the cores rotate the routes they are optimizing. For PSAD with mixing (PSAD-M), a group of cores optimize the same route using the same subset of pickup points. The cores within a group communicate in a self-adapted mixing period. While mixing, each core checks the similarity rate between its route with the current best in the group. If it is larger than a preset criterion, the current route will be replaced by the best.

We generate the framework for the published PSA methods and distinguish them by search space size, mixing period, and mixing pattern. Among these methods, we select the two most suitable methods, CDR and Lou'methods, and adapt them to this routing problem.

4 SPEEDUP

We use Intel Xeon E5-2690v3 CPU from Seawulf cluster at Stony Brook University for testing. We implement the parallel solutions using MPI. The problem settings for all the experiments are the same. The methods need to recommend 96 different routes for the drivers using 21,824 pick-up points in Beijing, China.

We compare the running time and the achieved MPTD between our proposed method, PSAD, to the other parallel methods with 96 cores. Our method spends 10.1 seconds to locate the solution while the other methods need more than one hour. On the other hand, our method achieves slightly better routes than those achieved by CDR and Lou'methods and much better than those by the parallel IBP and Enumerative process [1, 2, 6].

We also compare PSAD, CDR and Lou'methods with the same serial method [5]. By investigating the MPTDs obtained from the serial and parallel methods, we can calculate the lower and upper bound of the speedup. The speedup of PSAD for 96 computing cores is between 88 and 100 which is much higher than the CDR and Lou'methods whose speedups are below 4. The rotation operation makes PSAD to be more greedy than its serial version, which gives PSAD possibilities to achieve super speedup.

For PSAD, the parallel size cannot exceed the total route number for recommendation, 96. For further speedup, we use PSAD-M. For 384 cores, its speedup is between 174 and 180.

REFERENCES

[1] Yong Ge, Hui Xiong, Alexander Tuzhilin, Keli Xiao, Marco Gruteser, and Michael Pazzani. 2010. An energy-efficient mobile recommender system. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 899–908.

[2] Jianbin Huang, Xuejun Huangfu, Heli Sun, Hui Li, Peixiang Zhao, Hong Cheng, and Qinbao Song. 2015. Backward Path Growth for Efficient Mobile Sequential Recommendation. *IEEE Trans. Knowl. Data Eng.* 27, 1 (2015), 46–60.

[3] Rachel Sugar. [n. d.]. Uber and Lyft cars now outnumber yellow cabs in NYC 4 to 1. Who will sing for the yellow cab? Retrieved Oct 3, 2018 from https://ny.curbed.com/2017/1/17/14296892/yellow-taxi-nyc-uber-lyft-via-numbers

[4] Mohammed M Vazifeh, P Santi, G Resta, SH Strogatz, and C Ratti. 2018. Addressing the minimum fleet problem in on-demand urban mobility. *Nature* 557, 7706 (2018), 534.

[5] Zeyang Ye, Keli Xiao, Yong Ge, and Yuefan Deng. 2018. Applying Simulated Annealing and Parallel Computing to the Mobile Sequential Recommendation. *IEEE Transactions on Knowledge and Data Engineering* (2018).

[6] Zeyang Ye, Lihao Zhang, Keli Xiao, Wenjun Zhou, Yong Ge, and Yuefan Deng. 2018. Multi-User Mobile Sequential Recommendation: An Efficient Parallel Computing Paradigm. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2624–2633.