Deep Learning for Data-Driven Districting and Routing

A. Ferraz[a], Q. Cappart[b] and T. Vidal[a,c,*]

[a] Department of Computer Science, Pontifical Catholic University of Rio de Janeiro, Brazil
[b] CIRRELT, Department of Computer and Software Engineering,
École Polytechnique de Montréal, Canada
[c] CIRRELT & SCALE-AI Chair in Data-Driven Supply Chains, Department of Mathematics
and Industrial Engineering, École Polytechnique de Montréal, Canada
thibaut.vidal@polymtl.ca
* Corresponding author

Extended abstract submitted for presentation at the 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI)
June 19-25, 2022, Mauritius Island

January 15, 2022

Keywords: Districting, Routing, Graph Neural Network, Long-Term Cost Estimation

1 INTRODUCTION

Districting-and-routing is the process of partitioning a service region, represented as a collection of basic geographical units, into larger clusters called districts, and operating distinct routes within each district. This practice is ubiquitous in large-scale transportation and last-mile delivery systems for mail delivery, home care services, or maintenance services. A delivery policy in fixed districts has several interests: (1) allowing the separation and the aggregation of the requests in advance before all information is available, (2) reducing the complexity of the task thanks to the decomposition of the routing optimization process, (3) stimulating the familiarity of drivers and thus their efficiency within their respective geographical regions, and (4) increasing the satisfaction of customers thanks to a higher familiarity with their drivers.

Districting decisions are ubiquitous in large supply chains and linked with major financial and societal stakes. These decisions are strategic: they concern a few months or years, especially when dedicated facilities (e.g., warehouses) should be established. Simulating their impact, however, requires to evaluate the expected routing costs, which occur on a daily basis and vary with the demands. Because of these two different classes of decisions and planning horizons, the resulting problems pose considerable challenges (Drexl & Schneider, 2015).

Solution approaches for these problems can be generally divided into two classes, depending on how routing costs are evaluated. The first group of methods relies on continuous approximation models, e.g., Beardwood's TSP approximation formula, which estimates routing costs through n independently distributed points in a compact area of size A as $\alpha\sqrt{nA}$, where α is a constant (Franceschetti *et al.*, 2017). In contrast, the second group of methods explicitly solves routing sub-problems on a set of scenarios, a more accurate but time-consuming process. Yet, evaluation speed is critical when applying optimization techniques (e.g., local search) on the districting decisions, and it becomes prohibitively long to generate optimized routes for each scenario at each optimization step.

To fill this methodological gap, we capitalize upon the considerable progress in machine learning and deep neural networks to propose a solution approach which *learns* routing costs. More specifically, we train a *graph neural network* (GNN: Scarselli *et al.* 2008, Kipf & Welling

2016) to approximate the routing costs for a network of *contiguous geographical units*. We train the neural network on a set of examples that include the basic features of the geographical units in the city (density, area, shape, and adjacency matrix), as well as the routing costs estimated with Lin–Kernighan algorithm on a set of scenarios for different subsets of contiguous units.

2 METHODOLOGY

The problem is formalized as follows. A *region* $\mathcal{R}$ is divided into n *geographical units* $g^{\mathcal{R}}$, within which transportation operations will be independently performed over a long-term planning horizon. A *district* $\mathcal{D}^{\mathcal{R}} \subseteq \{1, \ldots, n\}$ of a region $\mathcal{R}$ is a subset of geographical units of $\mathcal{R}$. The *delivery cost* of a district is a function $C : \mathcal{D}^{\mathcal{R}} \to \mathbb{R}$, indicating the expected long-term cost of delivering customers in this district. Based on this context, the problem targeted by this paper is the design of districts such that the long-term operational cost is minimized. This is done under the assumption that the demand is modelled as a random distribution, independent and proportional to density of customers in an area. Besides, two constraints are considered: (1) the districts must be connected, and (2) the number of BUs inside each district is bounded to ensure a balance between vehicle workloads.

To search among possible districting solutions, any optimization method needs an oracle capable of accurately estimating the delivery cost C_d associated to a district d. Given the large number of districting-cost calculations in classical combinatorial optimization approaches, estimating C_d on a set of scenarios becomes a critical computational time bottleneck. In view of this, we propose to learn an approximate delivery cost function $\widehat{C}_d$ from historical data using a supervised learning approach together with a graph neural network. Figure 1 provides a high-level representation of the proposed network architecture. Our graph neural network is trained on a data set containing 10,000 districts, with their list of BUs and geographical characteristics. For each such district, the expected routing cost are calculated on a set of scenarios using the Lin-Kernighan heuristic of Helsgaun (2000).

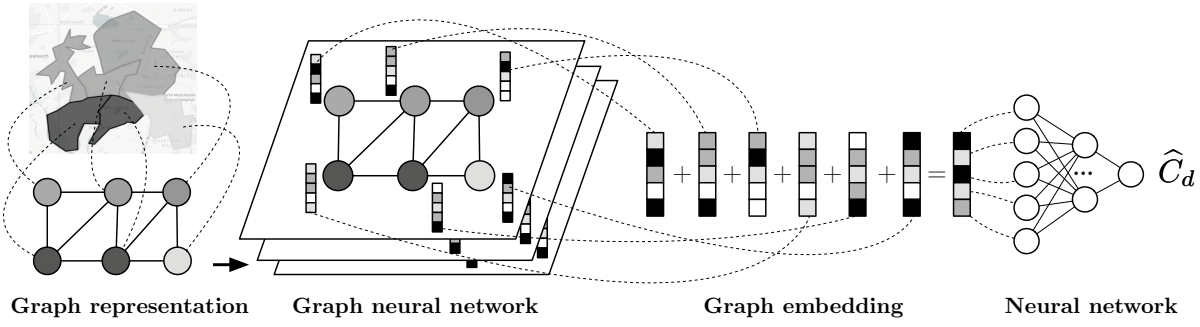

Figure 1 – *Neural architecture dedicated to estimate delivery costs* ($\widehat{C}^d$)

Once the neural network is trained, we rely on an iterative heuristic to find good districting plans. First we create an initial solution satisfying connectivity and balance requirements. This is equivalent to solving a balanced connected k-partition problem (Miyazawa *et al.*, 2020). We do so by representing it as a mixed integer program through a network flow formulation and solving it with CPLEX. Next, we improve the resulting solution using an Iterated Local Search (ILS) algorithm, using the routing-cost estimation oracle to evaluate whether moves are improving the current solution, and restricting the search to feasible solutions respecting the connectivity and balance constraints. We use Relocate and generalized Swap moves, whereas the perturbation operator consists in several combined random moves. The search method terminates after N_{IT} iterations of the local search and perturbation operator, and the best overall solution is returned.

To validate the performance of this heuristic, we also finally designed a set partitioning approach able to find optimal solutions for cases with small districts.

3 EXPERIMENTAL ANALYSES AND DISCUSSIONS

The goal of our experimental analyses is twofold: i) evaluating the accuracy of our GNN oracle for delivery-cost estimations, and ii) analyzing the impact of good cost estimations on the value of the districting solutions. As a baseline, we use a classic variation of the continuous approximation formula of Beardwood *et al.* (1959), which consists of a weighted sum of a) the distance between the depot and the closest point in the district, and of b) the term $\sqrt{nA}$ that estimates the internal routing costs. The weights are fine-tuned on a validation set prior to optimization. A similar formula been used in many previous studies (Franceschetti *et al.*, 2017). The resulting BD oracle can be used as a substitute for GNN in the districting solution method, and therefore permits to develop comparative analyses.

We rely on data sets built from five metropolitan areas in the United Kingdoms (Bristol, Manchester, Leeds, London and West-Midlands), with different geographical characteristics and varying population densities. The BUs correspond to the Middle Super Output Area (MSOA) in the regions, designed to have roughly a similar number of inhabitants (approximately 8,000). We assume that delivery requests are independent random events, with a frequency that is proportional to the number of inhabitants. We further create data sets considering different depot positions $\delta \in \{C, NE, NW, SE, SW\}$, number of BUs considered $n \in \{60, 90, 120\}$ and district-size targets $t \in \{3, 6, 12, 20, 30\}$ for the balancing constraints.

In a first experiment, we compare the quality of GNN and BD route-cost predictions. Table 1 reports the minimum squared error (MSE) evaluated in a validation set with the two different approaches, averaged over all considered depot-position configurations. Moreover, Figure 2 depicts the prediction performance of GNN and BD on the Bristol metropolitan area. We observe that GNN and BD predictions tend to be less accurate as the number of BUs in the districts increases. Nevertheless, the GNN approach is much more precise than BD, with an average MSE of 8.43 compared to 32.13.

	$n = 60$		$n = 90$		$n = 120$	
	GNN	**BD**	**GNN**	**BD**	**GNN**	**BD**
3	2.49	5.87	3.08	9.9	3.69	16.49
6	2.52	14.11	5.07	22.04	6.25	23.32
12	5.64	21.59	9.65	36.58	11.46	39.8
20	8.68	34.33	12.96	49.06	16.11	52.7
30	7.96	31.95	13.33	55.16	17.49	69.11

Table 1 – *Route-cost prediction performance*

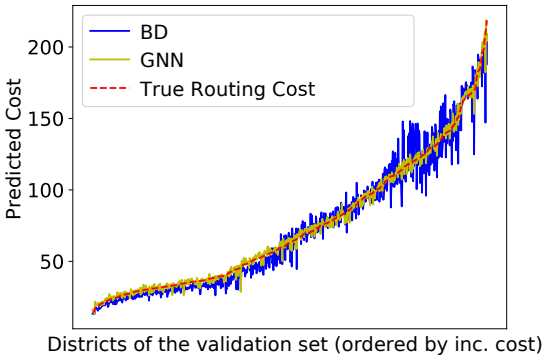

Figure 2 – *Prediction on Bristol-C-120-3*

After observing that the GNN generally provides more accurate routing cost estimates, we wish to evaluate the impact of a better oracle when searching for good districts. We therefore ran our ILS-based solution approach using either GNN or BD as an oracle, and measure the final solution quality using routing-cost simulation on a large number of scenarios. Table 2 compares the results of both approaches for different data set categories.

On average, the ILS using GNN achieved solutions with 6.45% lower long-term cost. Among important factors, we noticed that performance differences between BD and GNN are more significant when districts should be formed with a larger number of BUs (e.g., $t = 30$). This corresponds to a regime where GNN is known to be significantly better than BD (see Table 1). Moreover, the depot positioning also plays an important role. Improvement are more significant for GNN in cases where the depot is placed in the center. This is likely due to the fact that the share of distance to reach the districts is larger when the depot is placed in the corners, and this share of the cost is generally easier to estimate. Finally, we also observed that district compactness arose naturally as a by-product of the optimization oracle alone.

n	t	Depot Position					
		C	**NE**	**NW**	**SE**	**SW**	**Average**
60	3	1.82	0.86	0.83	0.31	0.76	0.92
	6	7.78	3.33	3.13	2.86	3.17	4.05
	12	12.54	7.19	3.69	4.54	5.96	6.78
	20	13.89	9.08	11.13	7.89	10.13	10.43
	30	12.48	9.30	8.17	7.45	8.28	9.14
90	3	1.90	-0.13	-0.35	-0.24	0.26	0.29
	6	6.41	1.95	1.93	1.82	2.65	2.95
	12	13.73	4.27	6.24	5.78	5.09	7.02
	20	14.62	11.89	8.15	7.76	9.22	10.33
	30	15.71	10.00	9.67	14.21	11.54	12.23
120	3	0.44	0.52	0.33	0.49	0.36	0.43
	6	3.96	1.91	2.47	2.95	3.20	2.90
	12	10.69	4.33	4.65	4.22	4.93	5.77
	20	17.35	9.05	9.65	5.41	13.86	11.06
	30	17.10	9.56	11.32	12.68	11.70	12.47
Average		10.03	5.54	5.40	5.21	6.07	6.45

Table 2 – *Districting-and-routing solution quality with BD or GNN*

4 CONCLUSIONS

As seen in this study, hybrid machine learning and optimization techniques can present notable advantages for strategic or stochastic problems in which it is challenging to efficiently evaluate long-term operational costs. Hereby, in the context of a districting-and-routing problem, the use of a GNN has permitted to obtain more accurate cost estimates as well significant long-term savings. The research perspectives are numerous and span three main directions. Firstly, more sophisticated network architectures could be developed to achieve better accuracy and generalization. Next, extensions of this solution paradigm could be devised for other classes of problems that include, e.g., production or facility location decisions. Lastly, other learning paradigms (e.g., reinforcement learning) may be exploited to allow an integration of the learned models in a mathematical programming approach (instead of an heuristic). These are all promising directions for future works.

References

Beardwood, J., Halton, J.H., & Hammersley, J.M. 1959. The shortest path through many points. *Mathematical Proceedings of the Cambridge Philosophical Society*, **55**(9), 299–327.

Drexl, M., & Schneider, M. 2015. A survey of variants and extensions of the location-routing problem. *European Journal of Operational Research*, **241**(2), 283–308.

Franceschetti, A., Jabali, O., & Laporte, G. 2017. Continuous approximation models in freight distribution management. *TOP*, **25**(3), 413–433.

Helsgaun, K. 2000. An effective implementation of the Lin-Kernighan traveling salesman heuristic. *European Journal of Operational Research*, **126**(1), 106–130.

Kipf, Thomas N, & Welling, Max. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Miyazawa, Flávio K., Moura, P.F.S., Ota, M.J., & Wakabayashi, Y. 2020. Cut and flow formulations for the Bbalanced connected k-partition problem. *Pages 128–139 of:* Baïou, M., Gendron, B., Günlük, O., & Mahjoub, A.R. (eds), *Combinatorial Optimization*. Cham: Springer International Publishing.

Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., & Monfardini, G. 2008. The graph neural network model. *IEEE Transactions on Neural Networks*, **20**(1), 61–80.