Point-to-point parcel delivery via clustering

Camille Gras, Van-Dat Cung, Nathalie Herr, Alantha Newman, Nicolas Teypaz

Point-to-point parcel delivery via clustering

Camille Gras[1 2], Van-Dat Cung[1], Nathalie Herr[2], Alantha Newman[1], Nicolas Teypaz[2]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP[†], G-SCOP, 38000 Grenoble, France

{van-dat.cung,camille.gras,alantha.newman}@grenoble-inp.fr

[2] ProbaYes, 38 330, Montbonnot, France

{camille.gras,nathalie.herr,nicolas.teypaz}@probayes.com

Keywords : *Operations research, Network design, Transport network, Optimization*

1 Introduction

We study a parcel delivery problem. Our exact problem is a generalization of a fixed destination point-to-point delivery problem known to be NP-hard [3]. Parcel delivery has been studied by several authors ; Zäpfel and Wasner [5] optimized the parcel delivery network in Austria (10 depots and one hub) using integer programming. They showed that a pure hub-and-spoke network is not optimal and studied the addition of direct paths between gathering points and dispersion points to a pure hub-and-spoke network. They concluded that a hybrid hub-and-spoke network is optimal in their case. Baumung and Gunduz used this type of network to consolidate residual volumes [1]. However in our case study the number of potential paths is too huge to be adressed by a MILP as shown in [2]. A similar situation was encountered for VRP, and Parmentier [4] discussed a clustering-based approach to address this issue. We will present an approach which combines consolidation of residual volumes and clustering ideas.

2 Definition : The parcel delivery problem

We address the problem of delivering parcels via truck in a national postal network composed of 18 origin sites and 136 destination sites. A parcel is a physical object that must be delivered from a specified starting site to a designated site within 36 hours. Each parcel is shipped in bulk in a container and each truck carries at most two containers. So a container carries parcels headed for one intermediate or final site as it cannot be partially unloaded. At an origin site, all parcels in a container can be *sorted* and each container can switch trucks. By sorting, we mean assigning parcels to containers and assigning containers to trucks. Each reassignment has a cost. Between the origin and the destination sites, a parcel can be reassigned (container and/or truck) at at most three sites. All pairs of sites are connected and there is a cost for routing a truck between two sites.

Our goal is to find paths for trucks in the transport network so that each parcel is delivered subject to the constraints we described above. The chosen paths and the routing have to be determined while minimizing the total transportation and sorting costs. Moreover, we need to "balance" the trucks, meaning that the number of outgoing and incoming trucks must be equal for each site in the course of a day.

Since a container can carry parcels with different origins and destinations, an optimal solution might not route a parcel on its shortest path. Thus our parcel delivery problem is not simply a shortest path problem. Moreover, we do not currently focus on the demand forecast problem, since it is a distinct problem and we assume that the demands (origin-destination-volume triples) are known for an average day. Finally, we note that the closest to our problem in the literature is called the fixed destination point-to-point-delivery problem [3].

[†]Institute of Engineering Univ. Grenoble Alpes

3 Clustering based solving methods

As the number of possible sorting operations for each parcel is bounded, we can set up a path-based MILP that is not too large in the sense that it can be given as an input for a standard LP-solver like CPLEX. However, it does not provide feasible solutions for a complete data instance in reasonable time [1]. But we can compute optimal solutions for smaller instances (less than 40 sites) in less than 10 minutes. This leads us to consider a divide-and-conquer approach, which we now describe.

Our algorithm consists of 3 parts : (1) a clustering, (2) a solving of the problem within each cluster and between clusters using MILP and (3) a global MILP to merge the solutions from part 2 in a cost-effective way. (1) The structure of the demand graph is such that there is no natural clustering in a graph theoretic sense (e.g., there are no sparse cuts). Moreover, if we have too many clusters, we will not be able to take full advantage of the benefits yielded by the consolidation of parcels. To find the best partition of sites, we compare various clustering algorithms (random, spectral clustering, hierarchical clustering, k-means) and apply them on the demand graph, the truck distance graph and the graph weighted by the number of possible paths between each pair of sites. We want well-balanced clusters because it will minimize the overall computation time. (2) In the second step, we solve the smaller versions of our problem on the clusters/regions using our MILP. We do the same for the problem between clusters. (3) Next, we solve a global MILP to merge the locally optimized solutions. And finally, we check the feasibilty of our solution. But the problem of routing between clusters still has too many variables and we need to investigate methods to decrease the size of the instance. For example, we can sparsify the path graph by pre-selecting some potential good paths.

Another approach is to first solve the inter-cluster problem and then the intra-cluster problems. (1) We use a clustering that has been created by experienced transport managers : A cluster around each sorting site contains all destination sites for which this sorting site is the nearest origin site. (2) We solve the inter-clusters MILP by sending aggregated demands between clusters. Then we use the MILP to optimize the routing between the sorting sites and the destination sites inside each cluster. (3) We merge the solutions of inter- and intra-clusters and then verify the feasibility of our merged solution.

Both these approaches give solutions for the complete instance in reasonable time, but they can remove some paths that an optimal solution would use. We will present experimental results obtained by implementing and analyzing these two methods.

References

[1] M. N. Baumung and H. I. Gündüz. Consolidation of residual volumes in a parcel service provider's long-haul transportation network. *Computational Logistics*, pages 437–450, 2015.

[2] C. Gras, V.-D. Cung, N. Herr, A. Newman, and N. Teypaz. Freight network design with logistics operation management. 20ème congrès de la ROADEF, 2019.

[3] C.-L. Li, S.T. McCormick, and D. Simchi-Levi. The point-to-point delivery and connection problems : complexity and algorithms. *Discrete Applied Mathematics*, 36(3), 1992.

[4] A. Parmentier. Du bon usage de l'apprentissage statistique en recherche opérationnelle. tutoriel au 20ème congrès de la ROADEF, 2019.

[5] G. Zäpfel and M. Wasner. Planning and optimization of hub-and-spoke transportation networks of cooperative third-party logistics providers. *International Journal of Production Economics*, 78(2) :207–220, 2002.

1. For 18 starting points, 136 endpoints, 2312 demands, 882 890 paths, 90 957 constraints, 28 million variables, the computation is not completed within 72 hours.