Vehicle routing optimization with relay: an arc-based column generation approach

Alexandre Jacquillat[a,b], Alexandria Schmid[b,*], Kai Wang[c]

[a] Sloan School of Management, MIT, Cambridge, MA, USA
[b] Operations Research Center, MIT, Cambridge, MA, USA
[c] Heinz College, Carnegie Mellon University, Pittsburgh, PA, USA
[*] 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 10, 2022

Keywords: relay logistics, physical internet, time-space network, column generation

1 INTRODUCTION

The logistics sector is a key enabler to growth and development. Decades of improvements in the design and operations of logistics systems have played an important role to support today's unprecedented growth of e-commerce as well as increasing delivery speeds. Yet, logistics providers face critical challenges to meet customer demand in increasingly interconnected networks, with limited capacity. In particular, spatial-temporal patterns routinely result in unnecessary vehicle miles traveled and empty backhauls. These inefficiencies have an important socio-economic toll, as exemplified by the low profit margins and high driver turnover faced by trucking operators. Moreover, the logistics sector contributes to nearly 20% of greenhouse gas emissions, creating intensive pressure to mitigate its environmental footprint (International Energy Agency, 2020).

On the bright side, freight transportation is witnessing several concomitant transformations. One area—at the core of this paper—lies in connected vehicles. By leveraging sensors and actuators, telematics provide near-real-time visibility into vehicle locations. Thanks to these data sources, digital platforms can provide decision-support systems to enhance freight operations. These advances have fueled new practices throughout supply chains based on collaboration, consolidation and cooperation which, collectively, can considerably improve the efficiency and sustainability of logistics operations (Savelsbergh & Van Woensel, 2016, Cleophas *et al.*, 2019).

One of these innovations is known as *relay-based logistics*. As opposed to point-to-point trucking operations, relay operations decompose each origin-destination trip into short segments traveled by separate drivers, using a network of "pit-stops". That is, a driver can move one order through one segment along one direction, drop the vehicle to another driver, and immediately travel along the opposite direction to serve another order. These operations rely heavily on connected vehicles and digitization to coordinate operations across the network of pit-stops. In the short term, relay logistics can improve drivers' lifestyles by allowing them to return home on a frequent basis. In the long term, they can play a catalyst role toward the "Physical Internet", by enabling modular logistics operations based on standardized and connected components (Montreuil, 2011).

To be successful, however, relay logistics require dedicated capabilities to coordinate operations across pit-stop networks. At the downstream level, relay operations raise the immediate question of which driver should cover which segment to move trucks and shipments across the network. But relay operations also require service providers to comprehensively re-optimize routing operations at the upstream level, in order to create relay opportunities and take full advantage of the coordination capabilities. In response, the goal of this paper is to develop new models and algorithms to optimize routing operations in relay-based logistics networks.

Specifically, we develop a new approach to optimize truck routes and driver assignments in order to best serve incoming orders in relay-based logistics. From a technical standpoint, we formulate a model that brings together several layers of time-space networks, with coupling constraints coordinating all entities. This approach, however, comes at the price of a large number of decision variables and constraints. We therefore develop a novel solution algorithm based on arc-based column generation, which starts from small time-space networks and expands them iteratively, until convergence. From a practical standpoint, we partner with Rivigo, a pioneer of relay-based operations in India. Using real-world data, we demonstrate the benefits of our modeling and algorithmic approach to maximize order throughput and minimize vehicle miles traveled in the Indian network. We also leverage the proposed modeling and algorithmic approach to evaluate the potential of relay-based operations in logistics. We detail the paper's contributions below.

2 MODEL FORMULATION

We develop an integer programming model that optimizes vehicle routes, driver assignments and order deliveries in relay-based logistics operations. It is formulated as a multi-commodity flow problem across four layers of time-space networks: (i) one for each customer order (ii) one for each driver (iii) one for empty trucks, and (iv) one for driver rebalancing. Accordingly, the model optimizes four sets of decisions: the flow of each order from an origin to a destination, the flow of each driver across the pit-stop network (while coming back home periodically), the flow of empty trucks, and the flow of rebalancing trips for drivers. As compared to traditional vehicle routing models (based on order-vehicle assignment), this problem is complicated by the coordination between driver operations and vehicle operations (based on order-truck-driver assignment).

The model balances two objectives, by minimizing miles traveled (a proxy for operating costs) and order delay (a proxy for level of service). Order-level constraints ensure that each order travels from origin to destination, maintain flow balance, and compute delays. Truck-level constraints maintain flow balance and apply initial truck availability. Driver-level constraints maintain flow balance and apply staffing rules. Finally, linking constraints ensure consistency between order movements, truck movements, and driver movements—namely, ensuring driver availability to cover truck-order assignments and empty trucks' trips.

By leveraging a multi-commodity flow structure, this model admits a tight linear programming relaxation. Its main complexity lies in its dimensionality. For drivers, staffing rules—in particular, each driver returning home regularly—lead to sparse time-space networks. For orders, however, time-space networks cannot be easily reduced because optimal solutions may involve detours to enable consolidation across the whole pit-stop network. One approach would be to proceed via heuristic arc reduction (e.g., a Boston–Miami order may not travel between Los Angeles and Seattle), but such heuristics can lead to sub-optimality or even infeasibility.

3 ARC-BASED COLUMN GENERATION

Instead, we propose a novel approach termed *arc-based column generation*. This approach starts with very sparse order-level time-space networks, and expands them by generating arcs iteratively until convergence. Note that the approach does not rely on a set partitioning formulation to generate *paths*, but instead uses the original formulation to generate *arcs* to build a small time-space network for each order. Specifically, the algorithm iterates between two modules:

- *Sparse time-space problem*: This problem generates a globally feasible solution by solving a variant of the model that restricts the set of arcs for each order. In the full problem, binary variables x_{ia} track whether order i travels through arc $a \in \mathcal{A}_i$ in its time-space network. The sparse time-space problem restricts these variables to a subset of arcs $\widehat{\mathcal{A}}_i \subseteq \mathcal{A}_i$.
- *Subproblem*: Using dual information from the sparse time-space problem, we compute the reduced cost associated with each order-arc pair (i, a). The subproblem seeks, for each order, a set of arcs that create a new path from origin to destination with minimal reduced cost. This is formulated as a shortest path problem on the full time-space network. We then add all corresponding arcs to the sparse time-space problem and proceed to Step 1.

As compared to "traditional" column generation methods, the subproblem does not seek a variable (or a column) with negative reduced cost, but rather a *set* of variables that create a path with negative total reduced cost. The algorithm terminates when there exists no such path for any order—although there may exist arc-based variables with negative reduced cost. We prove that this termination criterion guarantees optimality of the incumbent sparse time-space problem solution. As a result, the algorithm converges finitely to the optimal solution of the model's linear programming relaxation. We can embed it into a branch-and-price scheme to guarantee convergence to the solution of the full integer programming model.

Ultimately, this arc-based column generation algorithm provides a generalizable solution approach for time-space network optimization. As compared to heuristic arc reduction techniques, it yields an exact procedure to construct a reduced time-space network, without leading to potentially sub-optimal solutions or infeasible instances.

4 RESULTS

We implement our model and algorithm using real-world data from Rivigo's operations in 2019. We embed our arc-based column generation algorithm into a rolling implementation to evaluate its impact on relay-based logistics operations over a week of operations.

From a computational standpoint, we show that the proposed arc-based column generation algorithm converges effectively, in manageable computational times even for large-scale instances. Table 1 reports results of our algorithm against three benchmarks: direct implementation of the integer programming formulation, arc reduction heuristics, and "traditional" path-based column generation (based on an equivalent set partitioning formulation). The arc-based column generation algorithm yields near-optimal solutions by generating fewer arcs and terminating in shorter computational times than all benchmarks.

From a practical standpoint, Table 2 compares the solution to (i) a greedy, myopic heuristic, (ii) a sequential benchmark that first optimizes truck-order assignments (without considering relay) and then optimizes route-driver assignments (embedding relay dynamics ex post), and (iii) the historical routes implemented by Rivigo. Our solution reduces the costs by 10% as compared to Rivigo's historical operations, showing the benefits of optimization in complex relay-based logistics networks. Second, the orders-then-drivers heuristic increases costs by 20%. This underscores that relay-based operations do not simply require adjustments at the "downstream" level; instead, new relay capabilities require to re-optimize operations in order to create coordination opportunities between trucks and drivers. Furthermore, our solution achieves a Pareto improvement against all benchmarks—fewer miles traveled and lower order delay—therefore yielding win-win-win outcomes: better customer service, lower costs, and lower environmental footprint.

Table 1 – *Results of the arc-based column generation, against all computational benchmarks.*

Method	CPU Time (s)	Opt Gap	Arcs
Direct IP (full network)	23,470	0.00%	169,266
Arc reduction (aggressive)	5,665	3.51%	45,374
Arc reduction (moderate)	6,706	0.81%	93,583
Arc reduction (conservative)	18,697	0.01%	119,301
Path-based column generation	5,982	0.62%	-
Arc-based column generation	2,975	0.11%	8,165

Table 2 – *Solution from modeling and computational approach, against all practical benchmarks.*

Method	Cost	Total Miles	Empty Miles	Avg Delay	Orders Delivered
Proposed method	1,136,973	927,590	80,503	0.50	544
Greedy heuristic	1,371,241	1,005,927	155,561	0.94	462
	(+20.6%)	(+8.4%)	(+93.2%)	(+44.0%)	(-15.1%)
Orders-then-drivers	1,386,701	1,043,762	143,374	0.89	449
	(+22.0%)	(+12.5%)	(+78.1%)	(+39.0%)	(-17.5%)
Actual routes	1,255,678	971,771	98,200	0.75	469
	(+10.4%)	(+4.8%)	(+22.0%)	(+25.0%)	(-13.8%)

Finally, we compare relay operations to traditional point-to-point trucking in Table 3. For point-to-point, we introduce a parameter μ to incentivize drivers to come home: for small values of μ, drivers may be on the road indefinitely, whereas, for large values of μ, drivers will systematically return home even if this implies empty backhauling. First, note that the continuous movement of goods enabled by relay operations improves delay significantly when compared to point-to-point operations. Next, by design, relay operations allow drivers to return home much more frequently than point-to-point. Moreover, enabling drivers to come back home periodically in point-to-point operations (by increasing μ) has a significant cost in terms of empty miles. Ultimately, our modeling and computational approach again results in a Pareto improvement as compared to the no-relay benchmark: faster deliveries, better lifestyle for drivers, and fewer empty miles.

Table 3 – *Comparison of relay and non-relay operations on a medium-sized instance.*

Operations	μ	Total Miles	Empty Miles	Avg Delay	Orders Delivered	Nights Away
Point-to-point	0.01	491,932	58,779	1.77	536	3,248
Point-to-point	0.1	497,761	64,609	1.77	536	3,230
Point-to-point	1	545,860	112,707	1.77	536	2,018
Point-to-point	10	629,380	196,164	1.79	536	1,691
Point-to-point	100	648,140	214,469	1.79	533	1,915
Relay	-	516,715	91,143	0.55	613	963

References

Cleophas, Catherine, Cottrill, Caitlin, Ehmke, Jan Fabian, & Tierney, Kevin. 2019. Collaborative urban transportation: Recent advances in theory and practice. *European Journal of Operational Research*, **273**(3), 801–816.

International Energy Agency. 2020. *Energy Technology Perspectives 2020*. Tech. rept.

Montreuil, Benoit. 2011. Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, **3**(2), 71–87.

Savelsbergh, Martin, & Van Woensel, Tom. 2016. 50th anniversary invited article—city logistics: Challenges and opportunities. *Transportation Science*, **50**(2), 579–590.