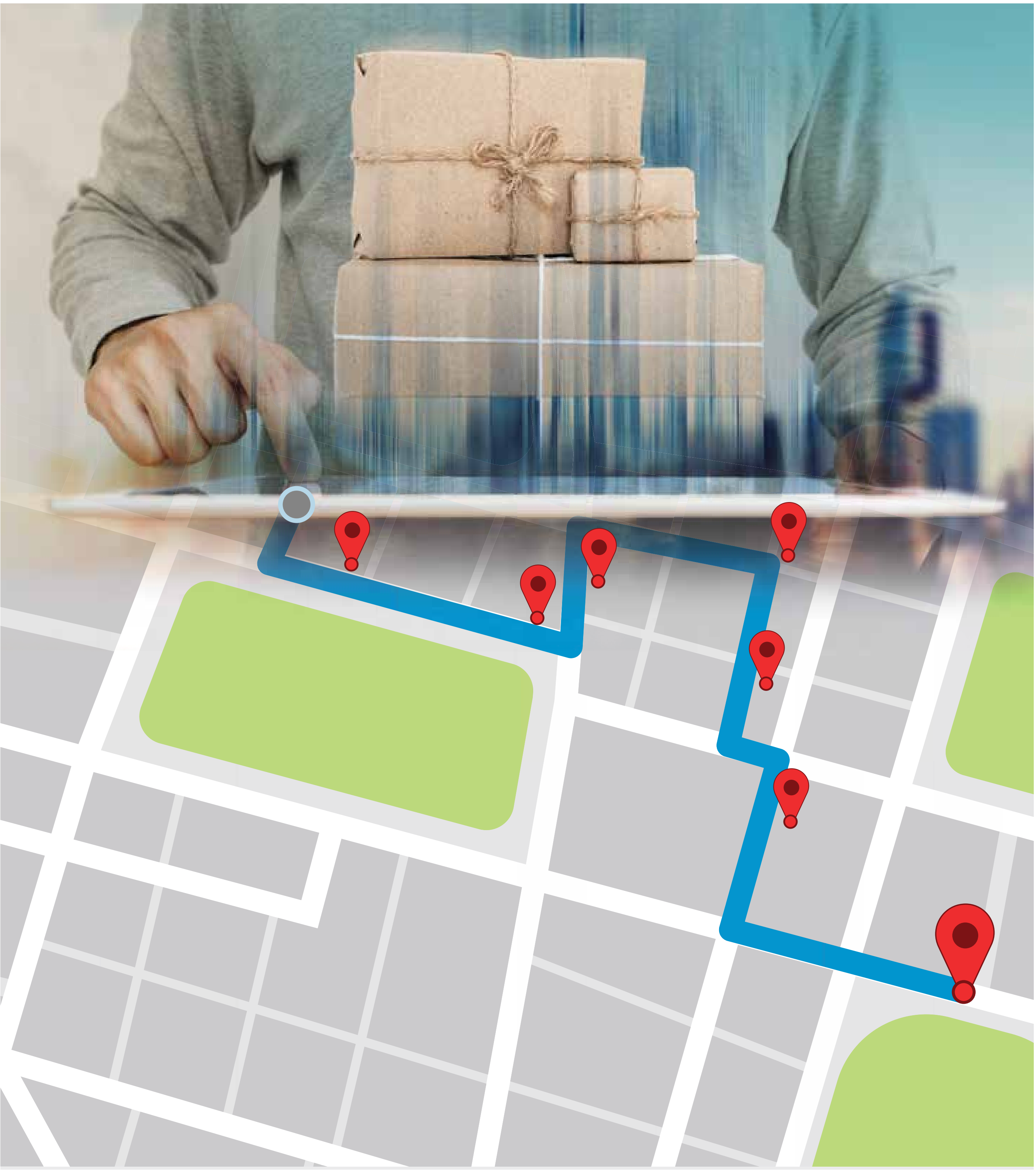
Last-mile Route Optimization

Whitepaper by Jitendra Kumar, Assistant Manager, Data Science | Ashutosh Vyas, Senior Data Science Manager - Mphasis NEXT Labs | Rahul Gupta, Data Science Manager – Mphasis NEXT Labs | Rohit Patel, AVP – Data Science and Quantum Computing - Mphasis NEXT Labs



Contents

1.
Introduction

Last-mile and first-mile logistics planning manages delivery and pickup of goods from a central location such as a hub, station or a depot to and from the end-customer location. In 2019, around 21.3 bn packages were delivered to end-customers worldwide and 15.5 bn packages were delivered in 2020 (as of Aug 2020). According to a Statista, 2018 report, USD 10.1 is the average delivery cost per package incurred by an organization. Of this, USD 8.1 per package is charged to customers, as against their willingness to pay around USD 1.4 per package. This shows a significant difference in current pricing and customer expectations.

Last-mile logistics contributes a massive 41% of the overall logistics cost. This increases to 53% when it comes to the overall shipping cost. Hence, optimizing the last-mile delivery operations can significantly reduce this difference. Freight forwarders such as FedEx have hubs and stations that deliver more than 20,000 packages per day. Planning optimal routes for a given fleet to deliver and pick up high-volume packages to end-customers becomes a non-trivial task requiring specialized optimization systems.

2.
Business Drivers and Challenges

Route planning and optimization is a complex process impacted by internal and external environmental factors. The major factors driving the complexity are the following:

- **Variability in shipment volume:** Daily shipment volume for delivery and pickup is highly impacted by customer behavior, temporal factors and operational complexity of the logistics service provider among others. Variability in shipment volume requires optimized and customized resource scheduling for daily operations.
- **Stress on customer-preferred time windows:** Customer-centric delivery models have gained momentum in the logistics industry with tighter delivery time windows. Pickup and delivery at the designated time windows by the customer not only makes it hard to minimize costs, but also demands flexible operating models to accommodate varying capacity requirements.
- **Diversified operating models:** Organizational business models guide decisions made regarding flexible vs fixed capacity fleet, drop and hook vs live load, full container load and less container load, etc. These choices come up with varied delivery operating models and associated costs.
- **Real-time traffic tracking and planning:** Real-time traffic monitoring and planning is a necessary component that captures deviations from the plan and re-allocates resources so that subsequent delivery planning is unaffected or minimally affected.
- **On-demand movements:** Tighter and faster responses to client requests is at the heart of business models. On-the-fly requests that route planning teams receive need to be accommodated in a cost-effective manner without affecting planned activities.

- **Legal constraints and commitments:** Legal constraints and regulations vary across regions. These regulations include driver working hours, shift assignment protocols, vehicle on-road restrictions, city entry restrictions, etc.
- **Fleet type and availability:** Fleet size, type and availability based on the maintenance schedule, ownership model, etc., can impact day-to-day operations.

3.
Optimization Objective & Constraints

The objective of last and first-mile route planning is to deliver and pick up goods and packages at an optimal operating cost; adhere to committed timelines, fulfill vehicle fleet size and capacity requirements; and drivers' working hours. Costs related to travel, resource assignment, vehicles, human resources and delays or penalties are generally considered to be constituents of the total cost, which needs to be minimized.

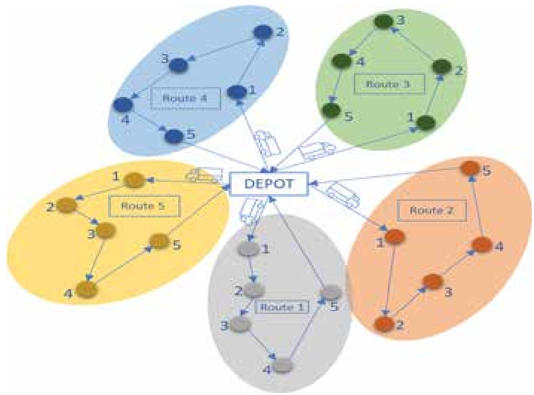
- **Travel cost** is proportional to the total distance traveled. The factors that influence travel cost include vehicle mileage, fuel cost, travel region and time, type of operating model, etc.
- **Resource assignment cost** is associated with assigning a resource (vehicle or human resource) to a movement. Nature of resource and resource utilization constraints will affect this cost.
- **Penalty cost** is related to non-adherence to customer Service Level Agreements (SLA). Delay in delivery is usually associated with a cost to the company, which can be included as part of route planning optimization. Sometimes, adherence to SLAs can be hard to achieve.

The main constraints for a route optimization problem are given below:

- **Fleet capacity management:** Mode of transport, number of vehicles and their capacity are significant factors that decide efficiency of route planning. Operational constraints related to material handling and operating business models also impact the optimal allocation process.
- **Driver work hour management:** The maximum allowed time of a driver on the road, shift regulations, movement restrictions are a few examples of legal constraints that must be adhered to.
- **Customer request management:** Adherence to customer requirements such as delivery and pickup time window is another important factor to be considered while route planning. Route planning needs to minimize the risks and costs associated with non-adherence to committed SLAs such as penalties, reputation loss, legal commitments, etc.

4.

Quantum Last-mile Route Optimizer



The route optimization engine has a two-fold task. Firstly, assignment of packages to vehicles within a given fleet and secondly, generate the order of deliveries for each vehicle in the fleet, based on pickup and deliveries schedule of a package from a single depot. The optimization engine minimizes the overall distance covered and hence the cost while managing the problem constraints.

Finding an optimal solution for route optimization is a compute-heavy and time-consuming process. As the size of the problem increases, there is an exponential increase in candidate solutions, and it becomes impossible to analytically solve the problem. Heuristics and metaheuristics-based optimization methods are used to find an approximate, but reasonable solution given the compute capacity and run-time constraints set by businesses. These methods use heuristics, metaheuristics and nature-inspired algorithms among others. Genetic algorithms, swarm optimization, ant colony optimization, etc., are some of the well-known techniques and tools that provide good approximate solutions.

In contrast to the classical methods mentioned above, in this paper, we address the last-mile optimization problem using quantum computing. Quantum computing systems make use of two models of computation, namely, Quantum gate-based computers and Quantum annealers.

Our primary digital model of classical computation is gate-based. We use these systems regularly for personal applications such as word processing, multimedia processing, networking and in many cases, managing industrial systems. These systems use elementary logic gates such as NOT, AND, OR, etc., to build digital circuits to solve problems. Quantum gate-based systems are the quantum counterparts of these classical digital circuits. They use reversible computation implemented via a different set of gates such as CNOT, TOFOLLI, HADAMARD, etc., to solve a computation problem.

Algorithms such as QAOA (Quantum Approximate Optimization Algorithm) utilize quantum properties such as superposition, entanglement and interference to solve optimization problems using gate model of quantum computation.

In the current solution though, we have used quantum annealers for quantum adiabatic optimization which uses the adiabatic theorem from quantum physics to minimize a function by interpolation between two Hamiltonians. Quantum annealers make use of a phenomenon observed in quantum systems called quantum tunneling for energy landscape exploration. This phenomenon can help quantum annealers realize a run-time and solution quality improvement on classical heuristics/metaheuristic counterparts such as simulated annealing, swarm optimization among others for certain types of problem energy landscapes.

4.1. Why Quantum for Last-mile Optimization?

Solving any optimization problem is a two-step process. The first step is to understand the problem and create a mathematical formulation of the problem. There are several ways to formulate a problem, such as Linear Programming (LP) formulation, Mixed Integer Linear Programming (MILP) formulation, non-linear formulation or quadratic formulation. The selection of formulation method is based on the convenience of formulating the problem, and the availability of algorithms, techniques and tools to solve the formulation.

The second step is to get the solution at optimal cost from the formulation. As large optimization problems cannot be solved analytically, and brute force methods can theoretically take exponentially a long time with respect to an increase in the input size, numerical methods are used to get approximate solutions. The current classical solutions are well developed through decades of research and can benefit from an increase in computing power and better algorithms.

Optimization problems can be converted into energy minimization problems, which in turn can be solved by quantum annealers. Quantum annealers follow nature-based optimization methodology and use energy encoding to map problems to hardware. It lets the system evolve with time, while controlling the rate of evolution, and given enough time, a system will reach the lowest energy point. Quantum annealers can solve optimization problems by improving solution quality and run-time using quantum mechanical properties such as quantum tunneling. Quantum annealers use Binary Quadratic Models (BQM) to formulate problem objectives and constraints, which in turn are converted into Quadratic Unconstrained Binary Optimization (QUBO) or equivalent ISING formulation in ferromagnetism.

In the context of last-mile route planning, the complexity of problems increases with increase in number of depots and customer locations, fleet size and business-specific constraints. Current classical route optimization systems are developed over decades and provide good approximate results for medium-sized problems. Additionally, many commercial logistics companies rely on homegrown algorithms, which optimize in parts to solve optimization problems. Such an approach can fall well short of the lowest cost of delivery. Quantum solvers such as hybrid classical-quantum solvers on quantum annealers and quantum-inspired classical optimization algorithms (QIO) can improve solution quality while reducing the run-time for certain types of energy landscapes representing a particular class of problems.

4.2. Solution Approach

As mentioned above, to use quantum annealers, the optimization problem must first be converted to a QUBO problem. The following requirements must be satisfied for each vehicle to find the minimum route cost for the problem.

- All customer locations must be covered
- All customer locations must be covered only once
- A customer location should be serviced by only one vehicle
- A vehicle cannot be loaded beyond its capacity
- A customer must be visited within the customer-provided time slots

The objective function of the problem is to minimize the total sum of travel costs for all vehicles.

QUBO formulation

For QUBO formulation, the constraint optimization problems are converted into unconstraint optimization problems using the penalty method. In the penalty method, constraints are added to the objective function by multiplying it with a penalty. The idea behind it is that if the solution fails to satisfy the constraints, then the penalty will be added to the total cost.

$$\begin{aligned}\text{Minimize } & -\sum_{k}^{K}\sum_{t}^{T}\sum_{i,j}^{n} d_{i,j} \times x_{k,t,i}\, x_{k,t+t_{ij},j} + A \times \sum_{k}^{K}\sum_{t}^{T}\sum_{i,j}^{n} x_{k,t,i}\, x_{k,t+t_{ij},j} \\ & + B \times \sum_{k}^{K}\sum_{i}^{n}\sum_{t_1,t_2}^{T} x_{k,t_1,i}\, x_{k,t_2,i} + C \times \sum_{i}^{n}\sum_{k_1,k_2}^{K}\sum_{t_1,t_2}^{T} x_{k_1,t_1,i}\, x_{k_2,t_2,i} \\ & + D \times \sum_{k}^{K}\sum_{i\,(i\neq 0)}^{n}\sum_{t_1,t_2\,(t_1<t_2)}^{T} x_{k,t_1,0}\, x_{k,t_2,i} + E \times \sum_{k}^{K}\left(\sum_{i}^{n}\sum_{t}^{T} L_i \times x_{k,t,i} - C_k\right)^2\end{aligned}$$

Here,

$x_{k,t,i}$: binary variable, =1 if customer "i" is visited at t^{th} time slot by k^{th} vehicle, =0 otherwise
n: number of customers
K: number of vehicles
T: number of time slots
$d_{i,j}$: distance between customer i and customer j
$t_{i,j}$: travel time between customer i and customer j
L_i : demand/load at customer i
C_k : capacity of vehicle k

A, B, C, D, E: large penalty for constraints

4.3. Experiment and Results

For this experiment, we have used 100 customer locations including a depot with a fleet size of 20 vehicles and a maximum capacity of 200 packages per vehicle. We have used three different solvers for this problem. All three solvers use Binary Quadratic Model (BQM) and hence Quadratic Unconstrained Binary Optimization (QUBO) formulation, to model objectives and constraints. The details are given below:

- **Qubovert simulated annealing solver** – An open-source simulated annealing algorithm for near-optimum global solution of a problem using probabilistic techniques.
- **D-Wave neal simulated annealing solver** – Simulated annealer from D-Wave.
- **D-Wave hybrid solver** – D-wave hybrid solvers implement state-of-the-art classical algorithms together with intelligent allocation of the Quantum Processing Unit (QPU) to parts of the problem where it benefits most.

	Qubovert Annealing	D-Wave Neal Simulated Annealing	D-Wave HSS
Dataset	Demand locations/points (including depot) = 101 Capacity of each truck (Q) = 200		
Number of Trucks	20		
Run-time (seconds)	Unable to converge	104*	33
Total Route Cost	Unable to converge	972*	1140

* Result with constraints violation.

As per the results, we see that D-wave hybrid solvers provide a feasible solution with an improved run-time for the given problem. In comparison, classical heuristics-based simulated annealer from D-Wave and Qubovert were unable to provide a solution or a feasible solution.

5.
Benefits/Advantages:

- **Fuel costs** comprise 25% of the total operational costs of a truck, according to American Transportation Research Institute (2019). With proper route optimization tools, a business could save up to 20% in mileage and improve order capacity by as much as 100% without increasing the fleet size.
- **Mphasis EON framework** for optimization on annealing-based systems allows for better model hyperparameter selection and experiment with what-if analysis. EON framework also manages problem formulation and model build lifecycle resulting in lower time to solution (approx. 40 times lesser).
- **Mphasis EON framework** allows for optimized what-if and solver re-runs. A focus on the re-usability of the previously built model obviates the need for model reformulation. A slight increase of 0.5% in total cost per movement is observed compared to a complete re-run of optimization, which saves 95% of time-to-solution and 98% in resource requirement.
- Based on our experiments, we have observed the improvements in the following indicative parameters on the quality of solution, when compared to classical metaheuristics solvers:
  - 30% less violations of legal and locational mandates
  - 15% reduction in delay costs as set by customer SLAs
  - A marginal 1.2% increase in travel cost
- Based on our experiments, we have observed that hybrid quantum solvers converge faster than their classical counterparts for certain problem types. Moreover, larger problem sizes, tend to provide more benefits leading to higher savings % (up to 230%).
- **EON-driven hybrid quantum optimization** on D-Wave annealing system delivers optimal and feasible solutions more frequently. The probability of getting feasible solutions as output has increased by 20% when compared to non-EON-driven quantum annealing.

6.
References

[1] https://blog.locus.sh/last-mile-delivery-costs-bringing-order-to-chaos/ - Last Mile Delivery Costs: Bringing Order to Chaos

[2] https://www.rcsgsolutions.com/post/outbound-shipping-and-the-cost-of-last-mile-delivery - Outbound Shipping and the cost of Last Mile Delivery

[3] https://arxiv.org/ftp/arxiv/papers/1404/1404.2465.pdf - Brief description on the state of the art of some local optimization methods: Quantum annealing

Authors

Jitendra Kumar

Assistant Manager, Data Science

Jitendra Kumar holds an M. Tech. degree from Indian Institute of Technology, Kanpur in Industrial & Management Engineering. His technical skills include Machine Learning, Deep Learning, Operations Research and Statistical Modeling. He has extensively worked in the field of predictive modeling, unsupervised ML techniques like clustering and optimization using annealing-based quantum systems.

Ashutosh Vyas

Senior Data Science Manager - Mphasis NEXT Labs

Ashutosh has 6+ years of experience in the data science domain. He has worked on multiple projects of pattern recognition, time series forecasting, regression modeling, NLP, classification and optimization in Life Science, Finance, FMCG and Media domains. He completed his MTech in 2015 from IIIT-B. He has expertise in Bayesian methods of Machine Learning and had been working in quantum ML and quantum optimization for the past 2 years and has developed multiple algorithms in image classification anomaly detection domain using quantum systems that leverage quantum gates and quantum annealer to process information and learn the patterns. At Mphasis, he works as a senior data science manager with an ethos of developing customer-centric and robust solutions.

Rahul Gupta

Data Science Manager – Mphasis NEXT Labs

Rahul Gupta holds an MBA degree from IIT, Kanpur with majors in Analytics and IT. Currently, he is working as Manager, Data Science. He has worked on several AI/ML projects with expertise in Image Analytics, Deep Learning, NLP and Quantum Machine Learning.

Rohit Patel

AVP – Data Science and Quantum Computing - Mphasis NEXT Labs

Rohit has 11+ years of experience in the IT industry with 6+ years of experience in Data Science. He holds a BTech degree in ECE and PGDM degree in Finance. He has been co-leading the Quantum Computing practice at Mphasis NEXT Labs for 2+ years. He has developed solutions in AI and Quantum Computing in the areas of Logistics, Life Sciences and Process Optimization, among others.

About Mphasis

Mphasis' purpose is to be the *"Driver in the Driverless Car"* for Global Enterprises by applying next-generation design, architecture and engineering services, to deliver scalable and sustainable software and technology solutions. Customer centricity is foundational to Mphasis, and is reflected in the Mphasis' Front2Back™ Transformation approach. Front2Back™ uses the exponential power of cloud and cognitive to provide hyper-personalized ($C = X2C^{2}_{TM} = 1$) digital experience to clients and their end customers. Mphasis' Service Transformation approach helps 'shrink the core' through the application of digital technologies across legacy environments within an enterprise, enabling businesses to stay ahead in a changing world. Mphasis' core reference architectures and tools, speed and innovation with domain expertise and specialization, combined with an integrated sustainability and purpose-led approach across its operations and solutions are key to building strong relationships with marquee clients. Click here to know more. (BSE: 526299; NSE: MPHASIS)

For more information, contact: marketinginfo.m@mphasis.com

USA
Mphasis Corporation
41 Madison Avenue
35th Floor, New York
New York 10010, USA
Tel: +1 (212) 686 6655

UK
Mphasis UK Limited
1 Ropemaker Street, London
EC2Y 9HT, United Kingdom
T : +44 020 7153 1327

INDIA
Mphasis Limited
Bagmane World Technology Center
Marathahalli Ring Road
Doddanakundhi Village, Mahadevapura
Bangalore 560 048, India
Tel.: +91 80 3352 5000

www.mphasis.com

NR 17/03/22 US LETTER BASIL7262