Delivery Optimization in Logistics Using Advanced Analytics and Interactive Visualization*

Leo Mršić
Algebra University College
Zagreb, Croatia
leo.mrsic@algebra.hr

Abstract*. Paper is based on research and PoC with goal to unlock business value in delivery optimization, using advanced analytics and interactive visualization. There are several common approaches including travelling salesman problem (TSP) that has various applications even in its purest formulation, such as planning or logistics. Knowledge base powered by large data set is described through development and final, interactive, form. Using common principles, paper describe business value that can be extracted using advanced analytics and interactive visualization, how to structure steps as best practice to achieve success and what additional benefits can be expected by supporting this approach. As part of research, various insights are extracted from PoC and presented in forms suitable for general understanding and future research.*

Keywords. delivery optimization, knowledge base, advanced visualization, big data

1 Introduction

Delivery optimization is among most popular research topics. Many practical routing problems involve finding paths or tours that traverse a set of paths in a graph. The aim of solving such problems is to find cycle with the least cost which covers all or a subset of arcs in a graph with or without constraints [Corberan and Laporte, 2014]. The Travelling salesman problem (TSP) and its variations are among most popular central research problems in path routing. The Travelling Salesman Problem (TSP) is about going to each point or node exactly once while returning to the original point or node (moving across in a cycle) and also taking the shortest route among all possible routes that fulfil certain criteria (if such a route exists). Finding such a cycle, perforce finding the possibly unique optimal cycle with the shortest distance, is not “NP-hard”. As popular variation, the Chinese Postman Problem (CPP) or Route Inspection Problem (RIP) is about visiting each route between nodes at least once while returning to the original node and taking the shortest route among all possible routes that fulfil other criteria (if such a route exists). A solution that takes each route exactly once is automatically optimal and called a Eulerian Cycle. Finding such a cycle is more “feasible” and can be used efficiently to create business value. This problem was first solved Mei-Ko Kuan, a Chinese mathematician, in 1962 [Kuan, 1962]. He was first one who consider this problem from the practical perspective of a postman picking up mail at the post office, delivering it along a set of streets, and returning to the post office. Since he must follow route at least once, the problem is referred to as "Chinese" postman problem to investigate how to cover every street along path and return to the post office under the least cost. In practice, there are a lot of applications of this technique such as road maintenance, waste collection, bus scheduling etc. The objective of these problems which applied the model based on route optimization is to find a route so that all of the edges of a given graph have been traversed at least once within minimum cost [Yilmaz at all, 2017].

2 Modern logistics

In the early days of computers, mathematicians hoped that someone would come up with a much better approach to large traveling salesman problems. Research community was looking for some algorithm that would allow computers to solve problem in a reasonable amount of time. But while computer scientists have made progress with specific scenarios, volume of data was rapidly progressing: identifying the shortest round-trip route for a 49-city map in the 1950s, a 2,392-city map in the 1980s and an 85,900-city map in 2006, putting pressure on algorithm that can efficiently solve every traveling salesman problem. According to a landmark paper published in 1972 [Karp, 1972], such a solution might not even be possible.

Modern technologies, especially big data tools and techniques, powered by mobile internet technologies made great success out of traditional services like ride-sharing systems such as UberPool [UberPool] and LyftLine [LyftLine]. Looking at model key business value, it is obvious that they were based on allowing multiple goals with similar itineraries and time schedules to share a delivering resource. By being able to analyse and monitor various data (often real-time) such solutions and models can significantly increase

*This paper is published and available in Croatian language at: http://ceciis.foi.hr delivery occupancy rate, alleviate traffic congestion, and reduce the energy consumption of urban commuting [Ma et all, 2013] [Braverman et all, 2017] [Biswas et all, 2017][Zhang et all, 2014].

Moving towards direction of modern logistics, during analysis process two key modules in the ride-sharing system will be borrowed": customer-vehicle matching and ride-sharing routing. The customer-vehicle matching module concerns with two issues: (a) how to group customers to create proper ridesharing opportunities, and (b) how to match the formed groups to vehicles. We will modify this approach, confront modern technologies with traditional approach and identify business value unlocked during research process.

2.1 Traditional KPI's in logistic

Delivery management as part of supply chain is a process that is often measured on its logistics costs. That approach is a root cause of various issues but also key point for business model efficiency. Various computations do not measure the supply chain or its performance and can include factors outside of the supply chain. In addition, the way accounting treats supply chain costs is dated. Fact is, delivery business is being transformed. The Amazon Effect has stimulated the beginning of a global supply chain revolution. It is moving beyond e-commerce/B2C and is crossing industries and markets. The underlying expectations are pushing delivery business transformation which include improved performance and new metrics. Supplementing the macro performance with segmented KPIs provides understanding and insight to what is happening on both centralize and "decentralized" views. It enables seeing underlying factors to the "corporate" measure.

As part of traditional metrics in postal delivery include process steps standardization to quantify most valuable asses: time.

Table 1. Receipt of the shipment: steps

Step	Type	Default time for processing
Start/sort	Package	00:00:10
Logging in	Package	00:00:30
Load	Package	00:00:20
Unload	Package	00:01:00
Document	Package	00:01:00
Total time: receipt		**00:03:00**

Table 2. Default time for different procedures (selection)

Procedure	Default time for processing
Collect post box	00:02:00
Telegram delivery	00:04:00
Package delivery	00:10:00
Letter delivery	00:00:20

Recent studies have been focusing on designing and optimizing the matching module [Braverman et all, 2017] [Biswas et all, 2017] [Zhang et all, 2014] so we will add postman delivery vehicle attributes as well.

Table 3. Delivery vehicle attributes

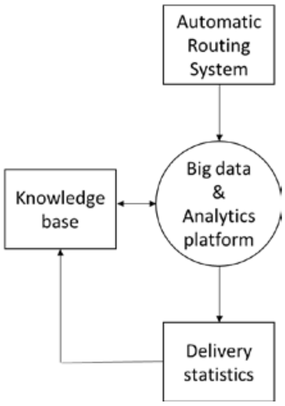
Route ID	Vehicle	Date/time	Description
RID-1	RN000	29.07. 14:11	Drive
RID-1	RN000	29.07. 14:22	Stop
RID-1	RN000	29.07. 15:11	Drive
RID-2	TZ100	29.07. 10:11	Drive

Modern routing platforms aims to determine a path between nodes in the formed groups while different choices of paths can lead to different cost or profit to the service providers. It is our aim to compare business value between scenario where delivery travel demands are given and delivery officer simply follow the path with minimum cost or maximum profit versus possibility for same system to be updated with knowledge base data, preferably in real time. Main difficulty of performing the matching and routing is that future demands are unknown when making decisions. Existing works outline two approaches to addressing this difficulty. The first one is to assume that all future travel demands are known at each matching decision timeframe, like offline approach [Biswas et all, 2017]. The second one is to assume zero knowledge of future demands [Zhang et all, 2014].

3 Delivery Optimization in Logistics Using Advanced Analytics and Interactive Visualization

With the recent technical and scientific advancements in analytics lead to exponentially growing sets of data. This generation of massive data along with the opportunities it provides for discovering new values and deriving new insights, and the various challenges it attempts to raise in terms of analysis and management, have created a new concept commonly referred to as Big Data [Borgi et all, 2017]. Logistics and transportation sectors are among the most ideally placed to take advantage of the analytical capabilities and the methodological advancements of Big Data technologies.

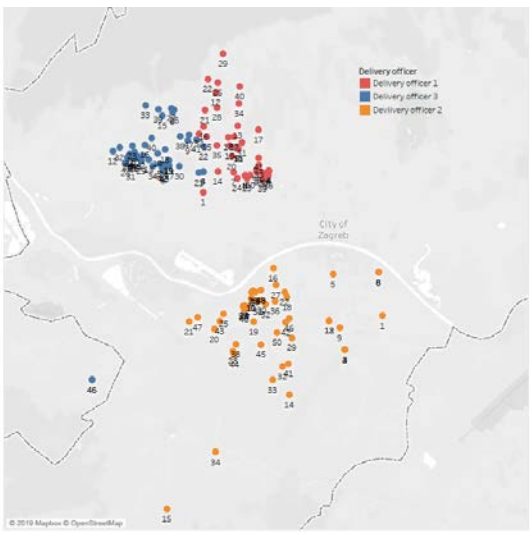
Figure 1. Big data analytics powered by knowledge base data feed



3.1 Delivery optimization

Our research data sample include three delivery officers working on hypothetic route. Route data was generated using City of Zagreb map, Croatia. Delivery officers were using various transportation support (motorcycle, delivery truck, bicycle) however each delivery officer was in charge of similar number of deliveries, approximately 50. We simulated urban area environment and look for business value that can be unlocked by developing knowledge base using all measurable data points following data privacy regulations.

Figure 2. Hypothetic delivery route



Delivery statistics shows several interesting points annotated with colour accordingly: red (unable to deliver), green (successful delivery), yellow (correct address, no one to deliver to), blue (unable to reach address), gray (no data).

An algorithm for finding an optimal route is:

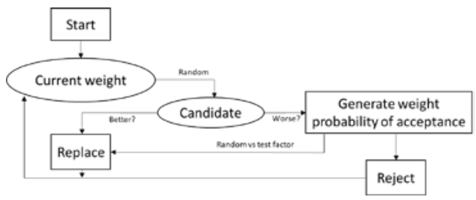
Step 1. List all odd vertices
Step 2. List all possible pairings of odd vertices
Step 3. For each pairing find the edges that connect the vertices with the minimum weight
Step 4. Find the pairings such that the sum of the weights is minimised
Step 5. On the original graph add the edges that have been found in Step 4.
Step 6. The length of an optimal route is the sum of all the edges added to the total found in Step 4.
Step 7. Select route corresponding to this minimum weight

Data modelling was dome using technique of combinatorial optimisation, simulated annealing. Process followed simple basic algorithm looking for maximums by raising weight (temperature) and then gradually reducing it, allowing local sections to grow outward. The algorithm randomly perturbs the original path to a decreasing extent according to a gradually decreasing logical weight (temperature). Spyder (Python 3.7) was used with anneal.py package[1].

To calculate weight between nodes we used formula below, while to find minimum to the objective function we adjust values of locationA and locationB.

$$objective = 0.2 + locationA_2^1 + locationB_2^2 - 0.1\ cox\ (6\pi x_1) - 0.1\cos(6\pi x_2)$$

Figure 3. Training algorithm



Training process for two nodes, looking for minimum weight between nodes (50 cycles, 50 trials per cycle):

Cycle: 0 with Weight: 2.8036732520571284
Cycle: 1 with Weight: 2.6391299733669826
Cycle: 2 with Weight: 2.484243487080245
…
Cycle: 34 with Weight: 0.3586429885950507
Cycle: 35 with Weight: 0.33759478221816164
Cycle: 36 with Weight: 0.3177818627582694
Cycle: 37 with Weight: 0.29913173312274866
…
Cycle: 48 with Weight: 0.28157615096203553
Cycle: 49 with Weight: 0.2650508789652898
Cycle: 50 with Weight: 0.249495449810821

Best objective: 0.008513963235742966

[1] http://apmonitor.com/me575/index.php/Main/SimulatedAnnealing

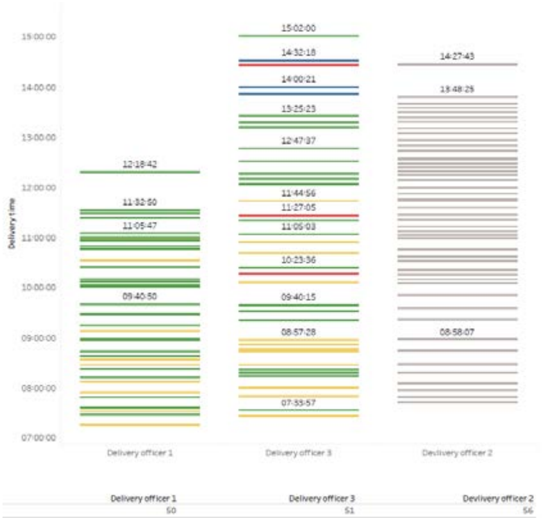
Results implemented into delivery sequence are shown in Table 3. below.

Table 3. Optimal delivery sequence

Delivery Officer	ID	Time of Delivery	Type	Lat/Long
DO1	107	14:20:44	Package	locationA
DO1	106	14:17:13	Letter	locationB
DO1	105	14:02:01	Package	locationC
…	…	…	…	…

In order to make delivery management more efficient, interactive visualization was made as shown in Figure 4. Delivery statistics on our simulation shows that there was difference in similar delivery scenario, when comparing time management as most valuable asset, by almost 25% between most efficient delivery officer and less efficient one.

Figure 4. Delivery statistics comparison



Using data retrieved from mobile devices, we are ready to serve knowledge database to see used for future improvements of the model (for example, if data shows delays in delivery for specific locations, we can inspect route).

3.2 Model improvement

Transportation is one of the leading industries to embrace technical innovations, simulation being number one among them. This game-changing technology is forming a new reality that is cost-efficient, eco-friendly, and customer-centred. Companies invest in simulation solutions (like High Performance Computing or HPC), and the tendency seems to be only picking up steam. The possibilities are endless as long as one understands how to find use case and move forward in creating business value. The data revolution caused by the arrival of HPC is not limited to enhancing business efficiency or improving service. It creates data flows that can be a precious and regular source of revenue for the businesses as well as for third parties. Many organizations are ready to purchase data that will help them understand their customers and boost their effectiveness. Therefore, for many companies the possibility to collect, treat and sell data is a possible monetization model. When this information has been accumulated and anonymized, companies can offer it packaged or raw, or sell it through various models. This allows them to improve manufacturing and service. But not just one company can take advantage of such data. Anonymized and proceeded it can be used by many other companies, which are ready to pay for such information. Customers use this anonymized data for a large variety of purposes, such as carbon footprint or costs reduction.

4 Conclusion

In our research we have focused on unlocking business value by creating data platform and challenge its value proposition with simple scenario. Our research shows that even on small scale business value can be created and can support decision for companies to invest in development of similar models.

We believe our research could be of great use for various logistics and transportation related computational scenarios. Paper is based on PoC that demonstrate usage of machine advanced analytics and interactive visualization. There are several common approaches including travelling salesman problem (TSP) that has various applications even in its purest formulation, such as planning or logistics. Knowledge base powered by large data set is described through development and final, interactive, form. Using common principles, paper confirm business value that can be extracted using machine learning and big data.

Best practice in how to achieve success and what additional benefits can be expected by supporting this approach are shown using various visualizations and presented in forms suitable for general understanding and future research.

References

A. Biswas, R. Gopalakrishnan, T. Tulabandhula, K. Mukherjee, A. Metrewar, and R. S. Thangaraj, “Profit optimization in commercial ridesharing,” in Proc. ACM AAMAS, 2017, pp. 1481–1483.

A. Braverman, J. Dai, X. Liu, and L. Ying, “Fluid-model-based car routing for modern ridesharing systems,” in Proc. ACM SIGMETRICS, 2017, pp. 11–12.

Corberan, A.; Laporte, G. Arc Routing Problems, Methods, and Applications. Society for Industrial and Applied Mathematics, USA, 2014.

D. Zhang, T. He, Y. Liu, S. Lin, and J. A. Stankovic, “A carpooling recommendation system for taxicab services,” IEEE Transactions on Emerging Topics in Computing, vol. 2, no. 3, pp. 254–266, 2014

Kuan, M. K. Graphic Programming Using Odd or Even Points. // Chinese Mathematics. 1962, pp. 237-277.

LyftLine. https://www.lyft.com/line.

M. Furuhata, M. Dessouky, F. Ordo´nez, M.-E. Brunet, X. Wang, and ˜S. Koenig, “Ridesharing: The state-of-the-art and future directions,” Transportation Research Part B: Methodological, vol. 57, pp. 28–46, 2013.

Mrsic, L. and Klepac, G and Kopal R (2017). A New Paradigm in Fraud Detection Modeling Using Predictive Models, Fuzzy Expert Systems, Social Network Analysis, and Unstructured Data, Computational Intelligence Applications in Business Intelligence and Big Data Analytics, Auerbach Publications, pp. 157-194

Mustafa Yılmaz, Merve Kayacı Çodur, Hamid Yılmaz (2017) Chinese postman problem approach for a large-scale conventional rail network in Turkey, Tehnički vjesnik 24, 5, pp. 1471-1477

Qiulin Lin, Lei Deng, Jingzhou Sun (2018), Minghua ChenOptimal Demand-Aware Ride-Sharing Routing, INFOCOM 2018

Richard M. Karp (1972). "Reducibility Among Combinatorial Problems", in R. E. Miller; J. W. Thatcher (eds.). Complexity of Computer Computations. New York: Plenum. pp. 85–103.

S. Ma, Y. Zheng, and O. Wolfson, “T-share: A large-scale dynamic taxi ridesharing service,” in Proc. IEEE ICDE, 2013, pp. 410–421.

Tawfik Borgi, Nesrine Zoghlami, Mourad Abed, Mohamed Saber Naceur (2017) Big Data for Operational Efficiency of Transportand Logistics: A Review, 6th IEEE International Conference on Advanced Logistics and Transport (ICALT)

UberPool. https://www.uber.com/ride/uberpool/.