Preference learning for solving constrained vehicle routing problems

Rocsildes Canoy and Tias Guns

Vrije Universiteit Brussel, Brussels, Belgium
{Rocsildes.Canoy,Tias.Guns}@vub.be

Abstract. The goal of this PhD research, which is done in collaboration with a Brussels-based logistics and sustainable transport solutions company, is to aid the company to better achieve its freight transport sustainability goals while respecting company constraints related to fleet size, vehicle capacities, etc. With the objective of learning the preferences of the route planners when choosing one option over another, and of effectively reusing all of the knowledge and effort that have been put into creating previous plans, we will investigate the use of machine learning and preference learning techniques over route plans. Our focus is on intelligent tools that learn from historical data, and can hence manage and recommend different or similar routes as used in the past. This is a novel research direction that will not only allow the company to innovate its freight transport planning process, but also encourage the active involvement of its employees and customers.

1 Research Progress and Motivation

The initial steps of the research included regular meetings with the CEO and representatives from the company, as well as an intensive review of the vast amount of literature that has been written on the different variants of the vehicle routing problem (VRP), the existing solution methods, and preference learning techniques. Weekly visits to the company depot were also made for data gathering, process shadowing, and brainstorming with the route planners.

From the knowledge acquired, we were able to formulate the research problem into a multi-objective, constrained, and capacitated vehicle routing problem. Route planning at the company is constrained by the limited number of vehicles, the capacity of each delivery vehicle, and the scheduling horizon within which all deliveries have to be made. The optimization objectives are in line with the company's goals of reducing operational costs, minimizing fuel consumption and carbon emissions, as well as maximizing driver familiarity with the routes.

We have learned that the daily plans are created in a route optimization software that is capable of producing plans that are optimal in terms of route length. It is the planners' usual practice, however, to either heavily modify the result given by the software, or simply to pull out, modify, and reuse an old plan that has been used and known to work in the past. By performing these modifications, the planners are essentially optimizing with their own set of objectives and personal preferences instead of relying on the optimal (distance-wise) plans given by the software.

In order to uncover the hidden preferences of the planners and to quantify the value that they attach to each sub-objective, we initially explored the idea of using state-of-the-art preference elicitation techniques. With each sub-objective given an initial weight estimate, one of the techniques, called the Preference Perceptron algorithm, solves the optimization problem and interacts with the user at each iteration, then uses the solution to update the current set of weights until either the user is satisfied or the process terminates after a fixed number of iterations. In either case, the learning algorithm yields a better estimate of the user weights, revealing the hidden preferences of the user. Using actual historical data obtained from the company, we implemented the algorithm in Python and performed initial experiments. Results, however, have not been especially convincing. Additional problem features or sub-objectives may possibly need to be considered. Also, to speed up the convergence of the algorithm, further fine-tuning or adjustments may be necessary, e.g., normalization to give a more or less identical scale to the sub-objective values.

2 Learning from Historical Solutions

Numerical analyses and data visualizations performed on the historical data have confirmed that the route planners often rely on past solutions in constructing the daily plans. This is consistent with the observations gathered during the company visits and has led us to investigate the possibility of using Markov decision process concepts in solving the route optimization problem. Given the current state (i.e., the present location of the vehicle), we make use of a probability matrix to determine the next state (where the vehicle goes to next). The probability matrix, which is based on historical data, is constructed in such a way that route arcs that have been frequently used in the past are assigned probabilities higher than those that were used less frequently. The optimization problem can now be solved using any constraint programming solver by maximizing the product of the probabilities of the arcs taken by the vehicles. This is a promising, novel approach to the vehicle routing problem. With the encouraging results that we have achieved from the preliminary experiments, we pursued the idea and eventually produced an **accepted paper** for CP 2019. Following is the abstract of our well-received paper, "**Vehicle routing by learning from historical solutions**:"

> *The goal of this paper is to investigate a decision support system for vehicle routing, where the routing engine learns from the* subjective *decisions that human planners have made in the past, rather than optimizing a distance-based* objective *criterion. This is an alternative to the practice of formulating a custom VRP for every company with its own routing requirements. Instead, we assume the presence of past vehicle routing solutions over similar sets of customers, and learn to make*

similar choices. The approach is based on the concept of learning a first-order Markov model, which corresponds to a probabilistic transition matrix, rather than a deterministic distance matrix. This nevertheless allows us to use existing arc routing VRP software in creating the actual route plans. For the learning, we explore different schemes to construct the probabilistic transition matrix. Our results on a use-case with a small transportation company show that our method is able to generate results that are close to the manually created solutions, without needing to characterize all constraints and sub-objectives explicitly. Even in the case of changes in the client sets, our method is able to find solutions that are closer to the actual route plans than when using distances, and hence, solutions that would require fewer manual changes to transform into the actual route plan.

3 Related Literature

The VRP becomes increasingly complex as additional sub-objectives and constraints are introduced. The inclusion of preferences, for example, necessitates the difficult, if not impossible, task of formalizing the route planners' knowledge and choice preferences explicitly in terms of constraints and weights. In most cases, it is much easier to get examples and historical solutions rather than to extract explicit decision rules from the planners, as observed by Potvin et al. in the case of vehicle dispatching [7]. One approach is to use learning techniques, particularly learning by examples, to reproduce the planners' decision behavior.

Learning from historical solutions has been investigated before within the context of constraint programming, e.g., in the paper of Beldiceanu and Simonis on constraint seeker [1] and model seeker [2], and Picard-Cantin et al. on learning constraint parameters from data, where a Markov chain is used, but for individual constraints [6]. In this respect, our goal is not to learn constraint instantiations, but to learn choice preferences, e.g., as part of the objective. Related to the latter is the work on Constructive Preference Elicitation [4], although that method actively queries the user, as does constraint acquisition [3].

Our motivation for Markov models is that they have been previously used in route prediction of individual vehicles. Krumm [5] has developed an algorithm for driver turn prediction using a Markov model. Trained from the driver's historical data, the model makes a probabilistic prediction based on a short sequence of just-driven road segments. Experimental results showed that by looking at the most recent 10 segments into the past, the model can effectively predict the next segment with about 90% accuracy. Ye et al. [9] introduced a route prediction method that can accurately predict an entire route early in the trip. The method is based on Hidden Markov Models (HMM) and also trained from the driver's past history. Another route prediction algorithm that predicts a driving route for a given pair of origin and destination was presented by Wang et al. [8]. Also based on the first-order Markov model, the algorithm uses a probability

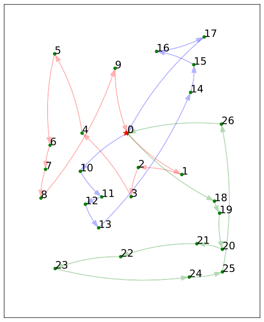

Fig. 1. Actual solution

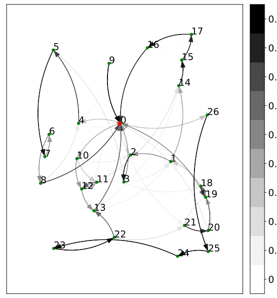

Fig. 2. Learned probabilities

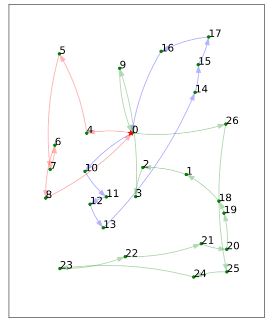

Fig. 3. Predicted solution

transition matrix that was constructed to represent the knowledge of the driver's preferred links and routes.

4 Results and Conclusion

The way the routes are predicted using a probability transition matrix can be observed from the visual example above. **Fig. 1** shows the actual route plan that we wish to reconstruct. The route plan is composed of three distinct routes, each originating from the depot, which is denoted by a red star. **Fig. 2** shows a visualization of the probability matrix learned from the historical solutions, with darker arcs indicating higher probabilities. The visualization shows a clear structure with distinct connections but also a higher variability in the denser regions and near the depot. **Fig. 3** shows our predicted solution, constructed with the probability matrix of **Fig. 2**. We observe that our solution captures key structural parts and makes trade-offs elsewhere to come up with a global solution. Furthermore, we see that the routes generally match and that it would require only a small amount of modifications to the predicted solution to obtain the actual solution.

In the paper, we presented an approach to solving the VRP which does not require explicit problem characterization. That is, we based our predictions on past solutions, thereby eliminating the need to explicitly define the problem's constraints and sub-objectives. Inspired by existing research on the application of Markov models to individual route prediction, we developed an approach that learns a probability transition matrix from previous solutions, to predict the routes for an entire fleet. This learned model can be transformed so that any CVRP solver (in our experiments, we used the CPLEX 12.8 solver) can be used to find the most likely routing. We showed how the structure of the solution can be learned, resulting in more accurate solutions than using traditional distance-based formulations. We observed that the algorithm is fast, and has the ability to learn the solution structure.

5 Future Work

Our first paper showed the potential of learning preferences in VRP from historical solutions. While results on the company data have been encouraging, validation on other real-life data will also be considered in the future. As we have so far only considered the case of the Capacitated Vehicle Routing Problem (CVRP), future work on the routing side will involve applications to richer VRP, e.g., problems involving time windows, multiple deliveries, etc. On the learning side, the use of higher-order Markov models or other probability estimation techniques in constructing the probability transition matrix will be investigated. Finally, extending the technique so that the user can be actively queried, and learned from, will be an interesting direction.

As the next step to the PhD research, we will focus our attention on the analysis of user preferences and user interactions with the optimization software, with the objective of incorporating active feedback to the system, rather than learning passively from historical instances. By observing how the planners adjust and modify the plans that have already been optimized by the system, we will be able to find out which information the planners base their decisions on. Our goal here is to develop a tool that can be integrated to the system and that can recommend actions to the user. Finally, once a working model is implemented and integrated into the company's daily operations, we will perform an analysis and assessment of the impact of the tool on all the involved stakeholders—planners, drivers, and customers.

References

1. Beldiceanu, N., Simonis, H.: A constraint seeker: Finding and ranking global constraints from examples. In: International Conference on Principles and Practice of Constraint Programming. pp. 12–26. Springer (2011)
2. Beldiceanu, N., Simonis, H.: A model seeker: Extracting global constraint models from positive examples. In: International Conference on Principles and Practice of Constraint Programming. pp. 141–157. Springer (2012)
3. Bessiere, C., Koriche, F., Lazaar, N., O'Sullivan, B.: Constraint acquisition. Artificial Intelligence **244**, 315–342 (2017)
4. Dragone, P., Teso, S., Passerini, A.: Constructive preference elicitation. Frontiers in Robotics and AI **4**, 71 (2018)
5. Krumm, J.: A markov model for driver turn prediction. In: SAE 2008 World Congress (April 2008), lloyd L. Withrow Distinguished Speaker Award
6. Picard-Cantin, É., Bouchard, M., Quimper, C.G., Sweeney, J.: Learning parameters for the sequence constraint from solutions. In: International Conference on Principles and Practice of Constraint Programming. pp. 405–420. Springer (2016)
7. Potvin, J.Y., Dufour, G., Rousseau, J.M.: Learning vehicle dispatching with linear programming models. Computers & operations research **20**(4), 371–380 (1993)
8. Wang, X., Ma, Y., Di, J., Murphey, Y.L., Qiu, S., Kristinsson, J., Meyer, J., Tseng, F., Feldkamp, T.: Building efficient probability transition matrix using machine learning from big data for personalized route prediction. Procedia Computer Science **53**, 284–291 (2015)

9. Ye, N., Wang, Z.q., Malekian, R., Lin, Q., Wang, R.c.: A method for driving route predictions based on hidden markov model. Mathematical Problems in Engineering **2015** (2015)