

MIT Open Access Articles

Predicting drivers' route trajectories in last-mile delivery using a pair-wise attention-based pointer neural network

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Mo, Baichuan, Wang, Qingyi, Guo, Xiaotong, Winkenbach, Matthias and Zhao, Jinhua. 2023. "Predicting drivers' route trajectories in last-mile delivery using a pair-wise attention-based pointer neural network." *Transportation Research Part E: Logistics and Transportation Review*, 175.

As Published: 10.1016/j.tre.2023.103168

Publisher: Elsevier BV

Persistent URL: <https://hdl.handle.net/1721.1/156448>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-ShareAlike



Predicting Drivers' Route Trajectories in Last-Mile Delivery Using A Pair-wise Attention-based Pointer Neural Network

Baichuan Mo^a, Qing Yi Wang^{a,*}, Xiaotong Guo^a, Matthias Winkenbach^b, Jinhua Zhao^c

^a*Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

^b*Center for Transportation and Logistics, Massachusetts Institute of Technology, Cambridge, MA 02139*

^c*Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 02139*

Abstract

In last-mile delivery, drivers frequently deviate from planned delivery routes because of their tacit knowledge of the road and curbside infrastructure, customer availability, and other characteristics of the respective service areas. Hence, the actual stop sequences chosen by an experienced human driver may be potentially preferable to the theoretical shortest-distance routing under real-life operational conditions. Thus, being able to predict the actual stop sequence that a human driver would follow can help to improve route planning in last-mile delivery. This paper proposes a pair-wise attention-based pointer neural network for this prediction task using drivers' historical delivery trajectory data. In addition to the commonly used encoder-decoder architecture for sequence-to-sequence prediction, we propose a new attention mechanism based on an alternative specific neural network to capture the local pair-wise information for each pair of stops. To further capture the global efficiency of the route, we propose a new iterative sequence generation algorithm that is used after model training to identify the first stop of a route that yields the lowest operational cost. Results from an extensive case study on real operational data from Amazon's last-mile delivery operations in the US show that our proposed method can significantly outperform traditional optimization-based approaches and other machine learning methods (such as the Long Short-Term Memory encoder-decoder and the original pointer network) in finding stop sequences that are closer to high-quality routes executed by experienced drivers in the field. Compared to benchmark models, the proposed model can increase the average prediction accuracy of the first four stops from around 0.2 to 0.312, and reduce the disparity between the predicted route and the actual route by around 15%.

Keywords: Route planning, Trajectory prediction, Sequence-to-sequence model, Last-mile delivery, Pointer network, Attention

1. Introduction

The optimal planning and efficient execution of last-mile delivery routes is becoming increasingly important for the business operations of many logistics service providers around the globe for a variety of reasons. E-commerce volumes are growing rapidly and make up a constantly growing share of overall retail sales. For instance, in the US, the share of e-commerce sales in total retail sales has grown from around 4% in 2010 to around 13% in 2021. Even by the end of 2019, i.e., before the outbreak of the COVID-19 pandemic,

*Corresponding author

it had reached 11% (US Census Bureau, 2021). Undoubtedly, the pandemic further accelerated the growth of e-commerce (postnord, 2021; McKinsey & Company, 2021). In the medium to long run, its growth will continue to be fueled by an ongoing trend towards further urbanization, which is particularly pronounced in developing and emerging economies (United Nations Department of Economic and Social Affairs, 2019). The share of the global population living in urban areas is currently projected to rise from around 55% in 2018 to around 68% by 2050. The associated increase in population density in most urban areas will likely lead to growing operational uncertainties for logistics service providers, as increasing congestion levels, less predictable travel times, and scarce curb space make efficient and reliable transport of goods into and out of urban markets increasingly challenging (Rose et al., 2016).

As a result of the continued boom of e-commerce and constantly growing cities, global parcel delivery volumes have been increasing rapidly in recent years and are expected to continue to do so. Across the 13 largest global markets, including the US, Brazil, and China, the volume of parcels delivered more than tripled from 43 billion in 2014 to 131 billion in 2020 (Pitney Bowes, 2020). At the same time, customer expectations towards last-mile logistics services are rising. For instance, there is a growing demand for shorter delivery lead times, including instant delivery services and same-day delivery, as well as customer-defined delivery preferences when it comes to the time and place of delivery (Lim and Winkenbach, 2019; Cortes and Suzuki, 2021; Snoeck and Winkenbach, 2021). The rapid growth and increasing operational complexity of urban parcel delivery operations also amplifies their negative externalities, including their contribution to greenhouse gas and other pollutant emissions, public health safety risks, as well as overall urban congestion and a corresponding decline in overall mobility and accessibility of cities (Jaller et al., 2013; World Economic Forum, 2020).

When applied to realistically sized instances of a last-mile delivery problem, solving the underlying traveling salesman problem (TSP) or vehicle routing problem (VRP) to (near) optimality becomes challenging, as both problem classes are known to be NP-hard. Traditional TSP and VRP formulations aim to minimize the total distance or duration of the route(s) required to serve a given set of delivery stops. The operations research literature has covered the TSP, VRP, and their many variants extensively, and in recent years important advances have been made with regards to solution quality and computational cost. However, in practice, many drivers, with their own tacit knowledge of delivery routes and service areas, divert from seemingly optimal routes for reasons that are difficult to encode in an optimization model directly. For example, experienced drivers may have a better understanding of which roads are hard to navigate, at which times traffic is likely to be bad, when and where they can easily find parking, and which stops can be conveniently served together. Therefore, compared to the theoretically optimal (i.e., distance or time minimizing) route, the *deviated actual route sequence* chosen by an experienced human driver is potentially preferable under real-life operational conditions.

An important challenge in today’s last-mile delivery route planning is therefore to leverage historical route execution data to propose planned route sequences that are close to the actual trajectories that would be executed by drivers, given the delivery requests and their characteristics. Note that, while distance and time-based route efficiency is still an important factor for planning route sequences, it is not the sole objective, as tacit driver knowledge is also incorporated in the proposed route sequences. Unlike a typical VRP in which the number of vehicles and their respective route sequences need to be determined simultaneously, in this study, we focus on solving a problem that is similar to a TSP at the individual vehicle level. That is, we aim to solve a stop sequence to serve a given set of delivery requests, and expect that the proposed stop sequence is as close to the actual trajectories that would be executed by drivers as possible.

To this end, we propose a pair-wise attention-based pointer neural network to predict the actual route sequence taken by delivery drivers using drivers’ historical delivery trajectory data. The proposed model follows a typical encoder-decoder architecture for the sequence-to-sequence prediction. However, unlike previous studies, we propose a new attention mechanism based on an alternative specific neural network (ASNN) to capture the local pair-wise information for each stop pair. To further capture the global efficiency of the route (i.e., its operational cost in terms of total distance or duration), after model training, we propose a new sequence generation algorithm that iterates over different first stops and selects the route with the lowest operational cost.

The main contribution of this paper is three-fold: First, we propose a new ASNN-based attention mechanism to capture the local information between pairs of stops (e.g., travel time, geographical relation), which can be well adapted to the original pointer network framework for sequence prediction. Second, we propose a new sequence generation algorithm that iterates over different first stops in the predicted route sequences and selects the lowest operational cost route. The intuition is that the stop-to-stop relationship (referred to as the *local view*) is easier to learn from data than the stop sequence of the route as a whole (referred to as the *global view*). Lastly, we apply our proposed method to a large set of routes executed by Amazon delivery drivers in the US. The results show that our proposed model can outperform traditional optimization-based approaches and other machine learning methods in finding stop sequences that are closer to high-quality routes executed by experienced drivers in the field.

The remainder of this paper is structured as follows. In Section 2 we define the problem setting under investigation in a more formal way. Section 3 then reviews previous studies in the literature related to this paper. Section 4 presents our methodology and elaborates on the detailed architecture of the proposed pair-wise attention-based pointer neural network. Section 5 presents the experimental setup and numerical results of our case study, applying our proposed method to real-world data made available by the Amazon Last-Mile Routing Research Challenge (Merchán et al., 2022; Winkenbach et al., 2021). Section 6 concludes this paper and discusses future research directions.

2. Problem Setting

In the last-mile delivery routing problem considered here, a set of stops $\mathcal{S} = \{s_1, \dots, s_n\}$ to be served by a given delivery vehicle is given to the route planner. The planner’s objective is to find the optimal stop sequence that has the minimal operational cost. In this case, we consider total cost as total travel time. The planner is given the expected operational cost (i.e., travel times) between all pairs of stops (s_i, s_j) . The theoretically optimal stop sequence, denoted by $(s_{(1)}^T, \dots, s_{(n)}^T)$, can be found by solving a TSP formulation. This stop sequence is referred to as the planned stop sequence. However, as discussed in Section 1, minimizing the theoretical operational cost (i.e., total travel time) of the route may not capture drivers’ tacit knowledge about the road network, infrastructure, and recipients. Therefore, the actual driver executed stop sequence $(s_{(1)}, \dots, s_{(n)})$ can be different from the planned route sequence. Note that here, $s_{(i)} \in \mathcal{S}$ denotes the i -th stop that is actually visited by the driver.

The objective of the model presented in this study is to predict the actual driver executed sequence $(s_{(1)}, \dots, s_{(n)})$ given a set of stops \mathcal{S} and the corresponding delivery requests and characteristics $X^{\mathcal{S}}$ (such as the number of packages, estimated service time for each package, geographical information for each stop, travel time between each stop pairs, etc.). All drivers are assumed to start their routes from a known depot $D^{\mathcal{S}}$ and return back to $D^{\mathcal{S}}$. Therefore, the complete trajectory should be a tour $(D^{\mathcal{S}}, s_{(1)}, \dots, s_{(n)}, D^{\mathcal{S}})$. For the convenience of model description, we ignore the depot station in the sequence.

Figure 1 provides a simple example for illustration. In this example, we are given four stops $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$ and a depot $D^{\mathcal{S}}$. The planned stop sequence for the driver is (s_4, s_1, s_2, s_3) , while the actual stop sequence executed by the driver is (s_4, s_2, s_1, s_3) . The proposed model aims to predict the actual sequence (s_4, s_2, s_1, s_3) given the depot location $D^{\mathcal{S}}$, the set of stops to be visited \mathcal{S} , and characteristics of the stops $X^{\mathcal{S}}$. This problem setup is inspired by the Amazon Last-Mile Routing Research Challenge (cf., [Winkenbach et al., 2021](#)). Note that this study only focuses on the stop sequence prediction. The routing between stops is not considered. It is assumed that the drivers always take the optimal route between stops, which is reflected by the travel time matrix between stops in our problem setup.

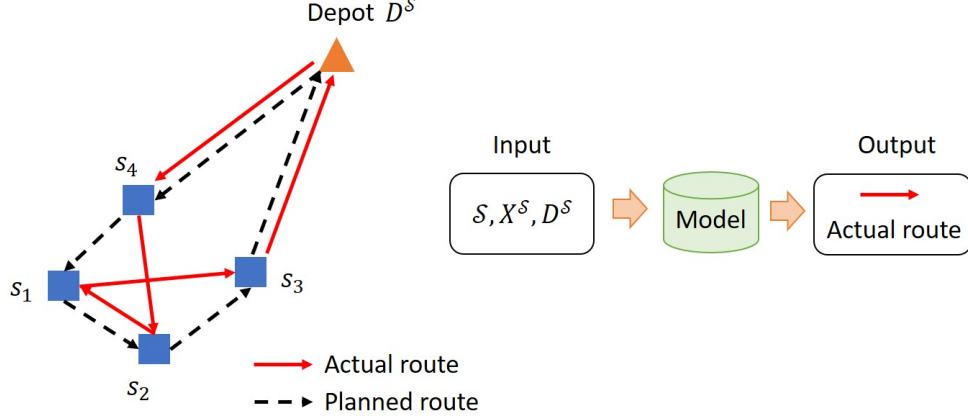


Figure 1: Illustrative example of the problem setting.

3. Literature Review

The problem setting defined in Section 2 involves both solving a cost-minimizing routing problem (i.e., the TSP) and capturing tacit driver knowledge to learn systematic deviation of drivers from the planned and theoretically optimal stop sequences. Therefore, we will first review the extant literature on the TSPs and its most relevant variants. We will then go through various machine learning approaches that have been proposed by the extant literature to generate sequences, with a section on methods specifically for solving the TSP. Note that although these machine learning approaches are used to solve the TSP instead of the actual routes taken by drivers, their architectures may be helpful to learn the actual route as well.

3.1. Travelling salesman problems

First, given the travel times between stops, a solution to the TSP, which finds the route with the minimum cost or distance (i.e., the planned route), can be a close approximation of the actual route. Since the drivers are paid for the number of packages delivered, all drivers' goal is to deliver the packages in the minimum amount of time. Most of the drivers do follow large parts of the planned routes.

The TSP is a well-known NP-hard problem that has been studied extensively over the last century, with a lot of books and review papers published on its history, formulations, solution approaches, and applications ([Applegate et al., 2006](#); [Matai et al., 2010](#); [Davendra and Bialic-Davendra, 2020](#)). An overview of the relevant TSP variants and solution approaches are presented below.

The basic setup of TSP has one traveler and requires the traveler to return to the starting point after visiting each node exactly once, and that the traveling cost matrix (represented by distance and/or time) is

symmetric (cost between i and j is the same with that between j and i). In most real-world applications, the basic setup needs to be modified. For example, the cost matrix, if represented by travel times, is likely asymmetric. This variant of TSP is thus named asymmetric TSP (ATSP) (Jonker and Volgenant, 1983). In some applications, the vehicle does not need to return to the original depot (Traub et al., 2021), or it can charge/refuel and potentially load additional delivery items at intermediate stops (Küçükoglu et al., 2019). In many last-mile delivery applications, some packages are time-sensitive, and therefore time window constraints to their delivery need to be considered in a so-called TSP with time windows (TSPTW) (da Silva and Urrutia, 2010; Mladenović et al., 2012). In large systems, there might be more than one salesman serving a set of stops, resulting in multiple traveling salesmen problems (MTSPs) (Cheikhrouhou and Khoufi, 2021).

Different variants of TSP further impose different constraints on the solution. While some problems can be reduced to the basic setup in the formulation stage, others require more versatile solution algorithms. In general, the solution approaches to the TSP can be divided into exact approaches and approximate approaches. Exact approaches include branch-and-cut (Yuan et al., 2020) and branch-and-bound (Salman et al., 2020). Since the TSP is a well-known NP-hard problem, exact approaches can only be applied on problems of smaller scale, or aid in heuristics to cut the solution space. Among approximate approaches, there are heuristics designed for the TSP specifically, as well as meta-heuristics that are generic and treat the problem like a blackbox. The most commonly used heuristics and meta-heuristics include nearest neighbor searches, local searches, simulated annealing, and genetic algorithms. A more comprehensive review of existing solution approaches can be found in Halim and Ismail (2017); Purkayastha et al. (2020). Despite the TSP being NP-hard, modern mixed-integer optimization solvers (e.g., Gurobi, CPLEX, or GLPK) can solve it efficiently for real-world instances by combining exact approaches with heuristics.

3.2. Sequence-to-sequence prediction using deep learning

The TSP and its variants are a viable option for sequence generation only when the objective is clearly defined. They fall short when the sequence generation problem does not have a well-defined cost-minimization objective. In a lot of applications, the rule of sequence generation cannot be simply defined and optimized.

A standard example for a sequence learning problem is machine translation, where a sequence of words in one language needs to be translated to another language. Another type of sequence learning is time series modeling, where a sequence of historical observations is given to predict future states of the system. In both cases, the primary modeling task is to learn the sequence generation rules. In recent years, deep learning has successfully achieved great performance in various settings of sequence generation. These models are often referred to as sequence-to-sequence (seq2seq) models.

seq2seq models often consist of an encoder and a decoder, where the encoder encodes the input sequence into a fixed-length vector representation, and the decoder generates a sequence based on the generated vector representation. Most encoder-decoder architectures adopt recurrent neural network (RNN) layers and its variants such as Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) and gated recurrent layers (GRU) (Cho et al., 2014) to learn long-range dependencies. Early works using LSTM alone were able to generate plausible texts (Graves, 2013) and translate between English and French (Sutskever et al., 2014) with long-range dependencies. Chung et al. (2014) demonstrate the superiority of GRU compared to LSTMs in music and speech signal modeling.

Attention-based mechanisms, first introduced by Bahdanau et al. (2015), have been shown to be a great addition since it allows the decoder to selectively attend to parts of the input sequence and relieves the encoder of the task of encoding all the information into a fixed-length vector representation. Most sequence generation

problems benefit from keeping track of long-range dependencies and global context while decoding. To address that, multi-level attention was proposed to capture the local and global dependency, and has shown to be effective in speech recognition (Chorowski et al., 2015), text generation (Liu et al., 2018), and machine translation tasks (Luong et al., 2015).

The encoder-decoder architecture combined with attention is very versatile, and it can be combined with other deep learning architectures to perform sequence learning in addition to language tasks. The LSTM and attention architecture is applied to semantic trajectory prediction (Karatzoglou et al., 2018), text summarization (Liang et al., 2020), demand modelling (Ren et al., 2020), and wind power forecasting (Zhang et al., 2020). When the goal is set to recover the original sequence, unsupervised learning of molecule embedding can be obtained for downstream classification tasks (Xu et al., 2017). When the spatial dimension is added, a convolutional neural network (CNN) layer can be added, and the dimension of the sequence generated can be expanded. For example, Wang et al. (2020a) predict a city’s crowd flow patterns, and Wu et al. (2020) generate 3D shapes via sequentially assembling different parts of the shapes.

While RNN-based architectures are still a widely adopted choice for seq2seq modeling, attention can also be used as a standalone mechanism for seq2seq translations independent of RNNs. The idea was proposed by Vaswani et al. (2017) in an architecture named transformer. Without recurrence, the network allows for significantly more parallelization, and is shown to achieve superior performance in experiments, and powered the popularity of transformer-based architectures in various sequence generation tasks (Huang et al., 2018; Lu et al., 2021). A separate line of work by Zhang et al. (2019) also demonstrated that a hierarchical CNN model with attention outperforms the traditional RNN-based models.

3.3. Using deep learning to generate TSP solutions

The above seq2seq translation mechanisms work well when the input data is naturally organized as a sequence, and the output sequence corresponds to the input sequence, such as in music and language. However, in our paper, the input is an unordered sequence, and the output has the same but re-ordered elements of the same input sequence. In this case, the concept of attention is helpful and has been successfully used to produce solutions to the TSP. The pointer network, proposed by Vinyals et al. (2015) and further developed in Vinyals et al. (2016), uses attention to select a member of the input sequence at each decoder step. While it is not required that the input sequence is ordered, an informative ordering could improve the performance (Vinyals et al., 2016).

While the original pointer network was solved as a classification problem and cross-entropy loss was used, it is not necessarily the most efficient choice. The cross-entropy loss only distinguishes between a correct prediction and an incorrect prediction. But in instances like routing, the distances between the predicted position and the correct position, as well as the ordering of subsequences, could incur different costs in practice. Further developments in solving TSP with machine learning methods involve reinforcement learning (RL), which enables the optimization of custom evaluation metrics (Bello et al., 2019; Kool et al., 2019; Ma et al., 2019; Liu et al., 2020). Joshi et al. (2019) compared the performance of RL and supervised learning (SL) on TSP solutions and found that SL and RL models achieve similar performance when the graphs are of similar sizes in training and testing, whereas RL models have better generalizability over variable graph sizes. However, RL models require significantly more data points and computational resources, which is not always feasible.

Although this seq2seq and attention framework has only been used to reproduce TSP solutions, it provides an opportunity to learn and incorporate additional information beyond the given travel times and potentially

learn individual differences when more information is given to the neural network. In this paper, we combine the ideas of seq2seq modeling and attention to predict the actual route executed by a driver.

4. Methodology

This section details the methodology proposed to address the problem. First, the high-level seq2seq modelling framework is introduced, followed by the explanation of the novel pair-wise attention and sequence generation and selection mechanism used within the modelling framework.

4.1. Sequence-to-sequence modeling framework

Let the input sequence be an arbitrarily-ordered sequence (s_1, \dots, s_n) . Denote the output sequence as $(\hat{s}_{(1)}, \dots, \hat{s}_{(n)})$. Let c_i indicate the “position index” of stop $\hat{s}_{(i)}$ with respect to the input sequence (where $c_i \in \{1, \dots, n\}$). For example, for input sequence (B, A, C) and output sequence (A, B, C) , we have $c_1 = 2$, $c_2 = 1$, $c_3 = 3$, which means the first output stop A is in the second position of the input sequence (B, A, C) and so on.

The seq2seq model computes the conditional probability $\mathbb{P}(c_1, \dots, c_n \mid \mathcal{S}; \theta)$ using a parametric neural network (e.g., recurrent neural network) with parameter θ , i.e.,

$$\mathbb{P}(c_1, \dots, c_n \mid \mathcal{S}, X^{\mathcal{S}}; \theta) = \mathbb{P}(c_1 \mid \mathcal{S}, X^{\mathcal{S}}; \theta) \cdot \prod_{i=2}^n \mathbb{P}(c_i \mid c_1, \dots, c_{i-1}, \mathcal{S}, X^{\mathcal{S}}; \theta) \quad (1)$$

The parameters of the model are learnt by empirical risk minimization (maximizing the conditional probabilities on the training set), i.e.,

$$\theta^* = \arg \max_{\theta} \sum_{\mathcal{S}} \mathbb{P}(c_1, \dots, c_n \mid \mathcal{S}, X^{\mathcal{S}}; \theta) \quad (2)$$

where the summation of \mathcal{S} is over all training routes. In the following section, we will elaborate how $\mathbb{P}(c_i \mid c_1, \dots, c_{i-1}, \mathcal{S}, X^{\mathcal{S}}; \theta)$ is calculated using the pair-wise attention-based pointer neural network.

4.2. Pair-wise attention-based pointer neural network

Figure 2 uses a four-stop example to illustrate the architecture of the proposed model. The whole model is based on the LSTM encoder and decoder structure. In particular, we use one LSTM (i.e., encoder) to read the input sequence, one time step at a time, to obtain a large fixed dimensional vector representation, and then to use another LSTM (i.e., decoder) to extract the output sequence. However, different from the typical seq2seq model, we borrow the idea of the pointer network (Vinyals et al., 2015) to add a pair-wise attention mechanism to predict the output sequence based on the attention mask over the input sequence. The pair-wise attention is calculated based on an ASNN which was previously used for travel mode prediction (Wang et al., 2020b). Model details will be shown in the following sections.

Intuitively, the LSTM encoder and decoder aim to capture the global view of the input information (i.e., overall sequence pattern) by embedding the input sequence to hidden vector representation. While the ASNN-based pair-wise attention aims to capture the local view (i.e., the relationship between two stops). Our experiments in Section 5 demonstrate the importance of both global and local views in the sequence prediction.

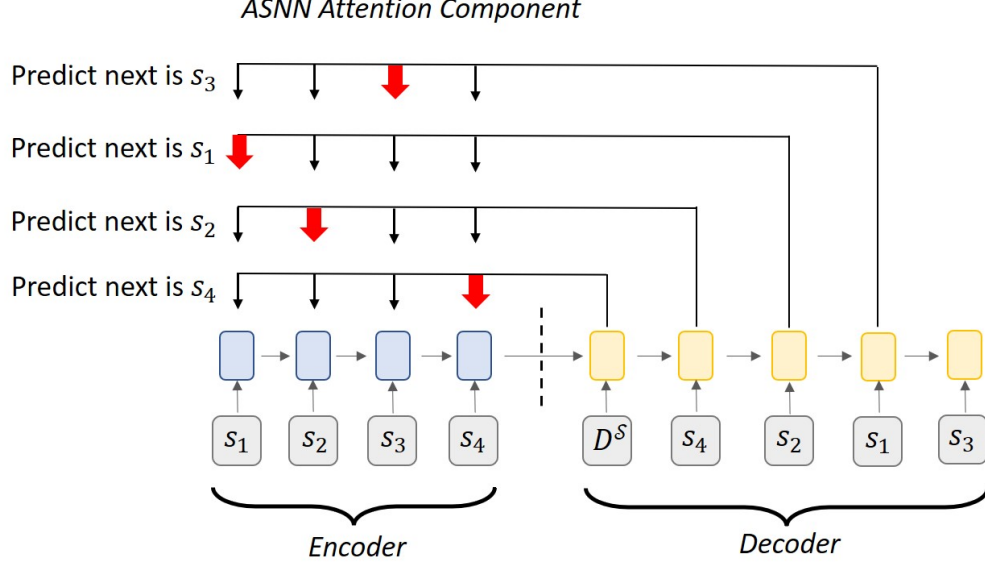


Figure 2: Overall architecture of the pair-wise attention-based pointer neural network (adapted from [Vinyals et al. \(2015\)](#))

4.2.1. LSTM encoder.

Given an arbitrary stop sequence (s_1, \dots, s_n) as the input, let $x_i \in \mathbb{R}^K$ be the features of stop s_i , where x_i may include the package information, the customer information, and the geographical information of the stop s_i . K is the number of features. The encoder computes a sequence of encoder output vectors (e_1, \dots, e_n) by iterating the following:

$$h_i^E, e_i = \text{LSTM}(x_i, h_{i-1}^E; \theta^E) \quad \forall i = 1, \dots, n \quad (3)$$

where $h_i^E \in \mathbb{R}^{K_h^E}$ is the encoder hidden vector with $h_0^E := 0$. $e_i \in \mathbb{R}^{K_e}$ is the encoder output vector. K_h^E and K_e are corresponding vector dimensions. θ^E is the learnable parameters in an encoder LSTM cell. The calculation details of an LSTM cell can be found in [Appendix A](#). The encoding process transforms a sequence of features (x_1, \dots, x_n) into a sequence of embedded representation (e_1, \dots, e_n) . And the hidden vector of the last time step (h_n^E) includes the global information of the whole sequence, which will be used for the LSTM decoder.

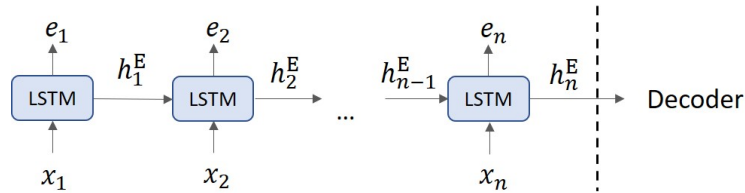


Figure 3: Illustration of LSTM encoder

4.2.2. LSTM decoder.

The role of a decoder in the traditional seq2seq model ([Figure 4](#)) is to predict a new sequence one time step at a time. However, in the pointer network structure with attention, the role of the decoder becomes producing

a vector to modulate the pair-wise attention over inputs. Denote the output sequence as $(\hat{s}_{(1)}, \dots, \hat{s}_{(n)})$. Let $x_{(i)}$ be the feature of stop $\hat{s}_{(i)}$.

At decoder step i , we have

$$h_{(i+1)}^D, d_{(i)} = \text{LSTM} \left(\begin{bmatrix} x_{(i)} \\ w_{(i)} \end{bmatrix}, h_{(i)}^D; \theta^D \right) \quad \forall i = 0, 1, \dots, n \quad (4)$$

where $h_{(i)}^D \in \mathbb{R}^{K_h^D}$ is the decoder hidden vector with $h_{(0)}^D = h_n^E$, $d_{(i)} \in \mathbb{R}^{K_d}$ is the decoder output vector, K_h^D and K_d are corresponding vector dimensions, and θ^D are learnable parameters of the decoder LSTM cell. Note that we set $x_{(0)} = x_D$ and $d_{(0)} = d_D$, representing the features and the decoder output of the depot, respectively. $w_{(i)}$ is the context vector calculated from the attention component, which will be explained in the next section.

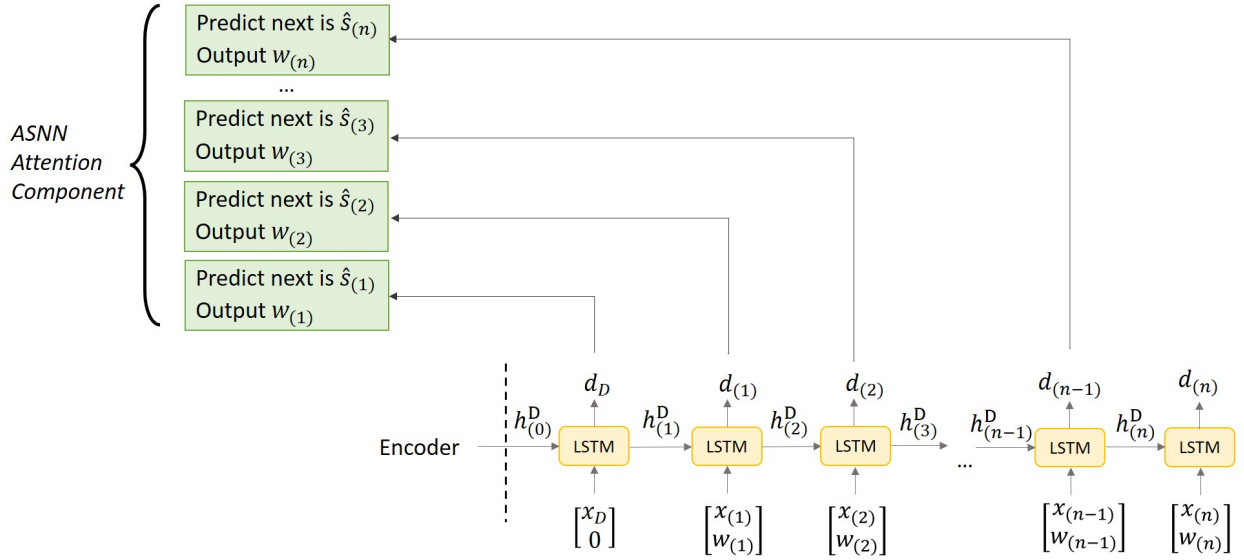


Figure 4: Illustration of LSTM decoder

4.2.3. ASNN-based pair-wise attention.

The pair-wise attention aims to aggregate the global and local information to predict the next stop. Specifically, at each decoder time step $i \in \{0, \dots, n\}$, we know that the last predicted stop is $\hat{s}_{(i)}$. To predict $\hat{s}_{(i+1)}$, we consider all candidate stops $s_j \in \mathcal{S}$, which is the set of all stops not yet visited. We want to evaluate how possible that s_j will be the next stop of $\hat{s}_{(i)}$. The information of the stop pair $\hat{s}_{(i)}$ and s_j can be represented by the following concatenated vector:

$$v_{(i)}^j = \text{concat}(z_{(i)}^j, \phi(x_{(i)}, x_j), d_{(i)}, e_j) \quad (5)$$

where $z_{(i)}^j$ is a vector of features associated with the stop pair (such as travel time from $\hat{s}_{(i)}$ to s_j), and $\phi(x_{(i)}, x_j)$ represents a feature processing function to extract the pair-wise information from $x_{(i)}$ and x_j . For example, $\phi(\cdot)$ may return geographical relationship between stops $\hat{s}_{(i)}$ and s_j , and it may also drop features not useful for the attention calculation. Intuitively, $z_{(i)}^j$ and $\phi(x_{(i)}, x_j)$ contains only local information of the stop pair, while $d_{(i)}$ and e_j contain the global information of the whole stop set and previously visited stops.

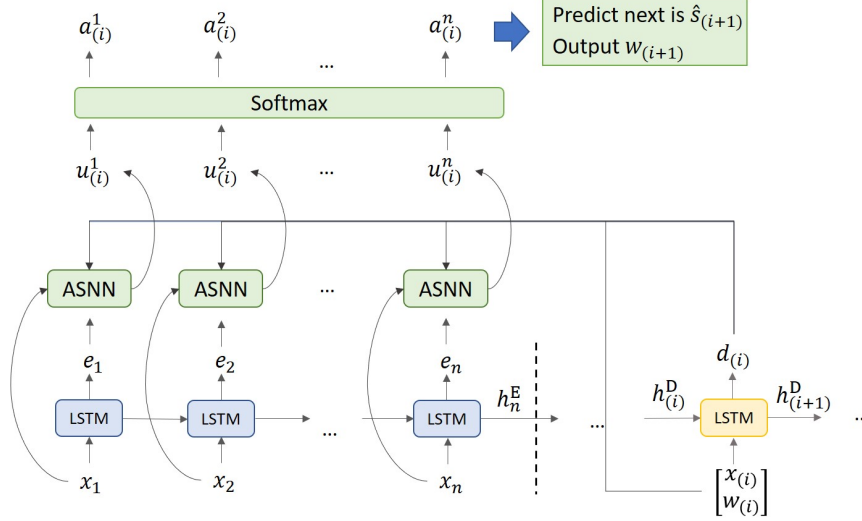


Figure 5: Illustration of ASNN-based pair-wise attention

Given the pair-wise information vector $v_{(i)}^j$, we can calculate the attention of stop $\hat{s}_{(i)}$ to stop s_j as:

$$u_{(i)}^j = \text{ASNN}(v_{(i)}^j; \theta^A) \quad \forall i, j = 1, \dots, n \quad (6)$$

$$a_{(i)}^j = \frac{\exp(u_{(i)}^j)}{\sum_{j'=1}^n \exp(u_{(i)}^{j'})} \quad \forall i, j = 1, \dots, n \quad (7)$$

where $a_{(i)}^j \in \mathbb{R}$ is attention of stop $\hat{s}_{(i)}$ to stop s_j . $\text{ASNN}(\cdot; \theta^A)$ is a multilayer perceptron (MLP) with the output dimension of one (i.e., $u_{(i)}^j \in \mathbb{R}$). θ^A are the learnable parameters of the ASNN. The name “alternative specific” is because the same parametric network will be applied on all alternative stops $s_j \in \mathcal{S}$ separately (Wang et al., 2020b). Finally, we calculate the conditional probability to make the prediction:

$$\mathbb{P}(c_{i+1} = j \mid c_1, \dots, c_i, \mathcal{S}, X^{\mathcal{S}}; \theta) = a_{(i)}^j \quad \forall i = 0, 1, \dots, n, j = 1, \dots, n \quad (8)$$

$$\hat{s}_{(i+1)} = \arg \max_{s_j \in \mathcal{S} \setminus \mathcal{S}_{(i)}^V} a_{(i)}^j \quad \forall i = 0, 1, \dots, n \quad (9)$$

where $\mathcal{S}_{(i)}^V = \{\hat{s}_{(1)}, \dots, \hat{s}_{(i)}\}$ is the set of stops that have been predicted (i.e., previously visited) until decoder step i . Eqs. 8 and 9 indicate that the predicted next stop at step i is the one with highest attention among all stops that have not been visited.

The pair-wise attention framework also leverages the attention information as the input for the next step. This was achieved by introducing the context vector (Bahdanau et al., 2015):

$$w_{(i)} = \sum_{j=1}^n a_{(i)}^j \cdot e_j \quad (10)$$

The context vector is a weighted sum of all the encoder output vectors with attention as the weights. As the attention provides the emphasis for stop prediction, $w_{(i)}$ helps to incorporate the encoded representation of the last predicted stop for the next stop prediction. The inputs for the next LSTM cell thus will be the

concatenation of the stop features and $w_{(i)}$, i.e., $\begin{bmatrix} x_{(i)} \\ w_{(i)} \end{bmatrix}$.

It is worth noting that, the specific architecture of ASNN($\cdot; \theta^A$) can be flexible depending on the input pair-wise information. For example, if the information includes images or networks, convolutional neural network or graph convolutional networks can be used for better extract features. In this study, we use the MLP for simplification as it already outperforms benchmark models. The key idea is of the ASNN is to share the same trainable parameter θ^A for all stop pairs so as to better capture various pair-wise information in the training process.

4.3. Sequence generation and selection

During inference, given a stop set \mathcal{S} , the trained model with learned parameters θ^* are used to generate the sequence. Typically, in the seq2seq modeling framework, the final output sequence is selected as the one with the highest probability, i.e.,

$$(s_{j_1^*}, \dots, s_{j_n^*}), \text{ where } j_1^*, \dots, j_n^* = \arg \max_{j_1, \dots, j_n \in \mathcal{C}^{\mathcal{S}}} \mathbb{P}(c_1 = j_1, \dots, c_n = j_n \mid \mathcal{S}, X^{\mathcal{S}}; \theta^*) \quad (11)$$

where $\mathcal{C}^{\mathcal{S}} = \{\text{All permutations of } \{1, \dots, n\}\}$

Finding this optimal sequence is computationally impractical because of the combinatorial number of possible output sequences. And so it is usually done with the greedy algorithm (i.e., always select the most possible next stop) or the beam search procedure (i.e., find the best possible sequence among a set of generated sequences given a beam size). However, in this study, we observe that the first predicted stop $\hat{s}_{(1)}$ is critical for the quality of the generated sequence. The reason may be that the local relationship between a stop pair (i.e., given the last stop to predict the next one) is easier to learn than the global relationship (i.e., predict the whole sequence). Hence, in this study, we first generate sequences using the greedy algorithm with different initial stops, and select the one with the lowest operational cost. The intuition behind this process is that, once the first stop is given, the model can follow the learned pair-wise relationship to generate the sequence with relatively high accuracy. For all the generated sequences with different first stops, the one with the lowest operation cost captures the global view of the sequence's quality. Therefore, the final sequence generation and selection algorithm is as follows:

Algorithm 1 Sequence generation

Input: Trained model, \mathcal{S}

Output: Predicted stop sequence

- 1: **for** s in \mathcal{S} **do**
 - 2: Let the first predicted stop be $\hat{s}_{(1)} = s$
 - 3: Predict the following stop sequence $(\hat{s}_{(2)}, \dots, \hat{s}_{(n)})$ using the greedy algorithm. Denote the predicted sequence as P_s .
 - 4: Calculate the total operation cost of the whole sequence (including depot), denoted as OC_s .
- return** P_{s^*} where $s^* = \arg \min_{s \in \mathcal{S}} OC_s$
-

5. Case Study

5.1. Dataset

The data used in our case study was made available as part of the Amazon Last Mile Routing Research Challenge (Merchán et al., 2022). The dataset contains a total of 6,112 actual Amazon driver trajectories

for the last-mile delivery from 5 major cities in the US: Austin, Boston, Chicago, Los Angeles, and Seattle. Each route consists of a sequence of stops. Each stop represents the actual parking location of the driver, and the package information (package numbers, package size, and planned service time) associated with each stop is given. The stops are characterized by their latitudes and longitudes, and expected travel time between stops are known.

Figure 6 shows the distribution of the number of stops per route and an example route. Most routes have around 120 to 180 stops, and the maximum observed number of stops is around 250. Figure 6b shows an example of an actual driver trajectory in Boston. Since the depot is far from the delivery stops, we attach the complete route (with the depot indicated by a red dot) at the bottom left of the figure, while the main plot only shows the delivery stops.

In this data set, each stop is associated with a zone ID (indicated by different colors in Figure 6b). When Amazon generates planned routes for drivers, they usually expect drivers to finish the delivery for one zone first, then go to another zone. And the actual driver trajectories also follow this pattern as shown in Figure 6b (but the actual zone sequence may be different from the planned one). Therefore, in this study, we focus on the problem of *zone sequence prediction*. That is, s_i in the case study section now represents a specific zone, \mathcal{S} represents the set of zones, and $X^{\mathcal{S}}$ represents zone features. This transformation does not affect the model structure proposed in Section 4. The only difference is that the new problem has a relatively smaller scale compared to the *stop sequence prediction* because the number of zones in a route is smaller than that of stops. The zone-to-zone travel time is calculated as the average travel time of all stop pairs between the two zones. Figure 7 presents an illustrative example of the relationship between zone and stop sequences. As the dataset does not contain the original planned sequence, we assume the planned zone sequence is the one with the lowest total travel time (generated by a TSP solver, (s_1^T, \dots, s_n^T)). After generating the zone sequence, we can restore the whole stop sequence by assuming that drivers within a specific zone follow an optimal TSP tour. Details of the zone sequence to stop sequence generation can be found in Appendix B.

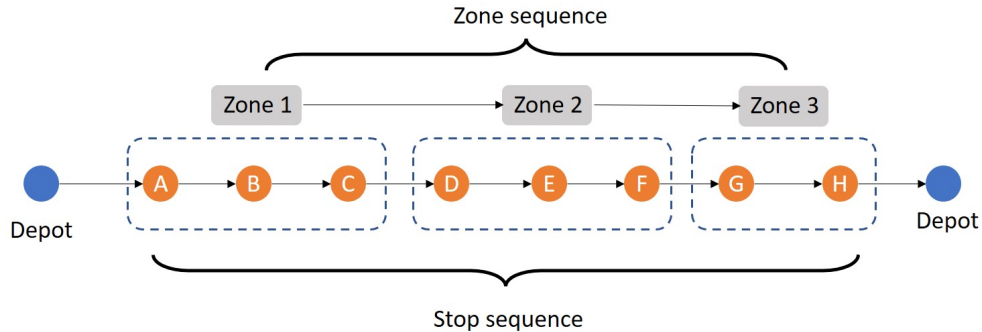
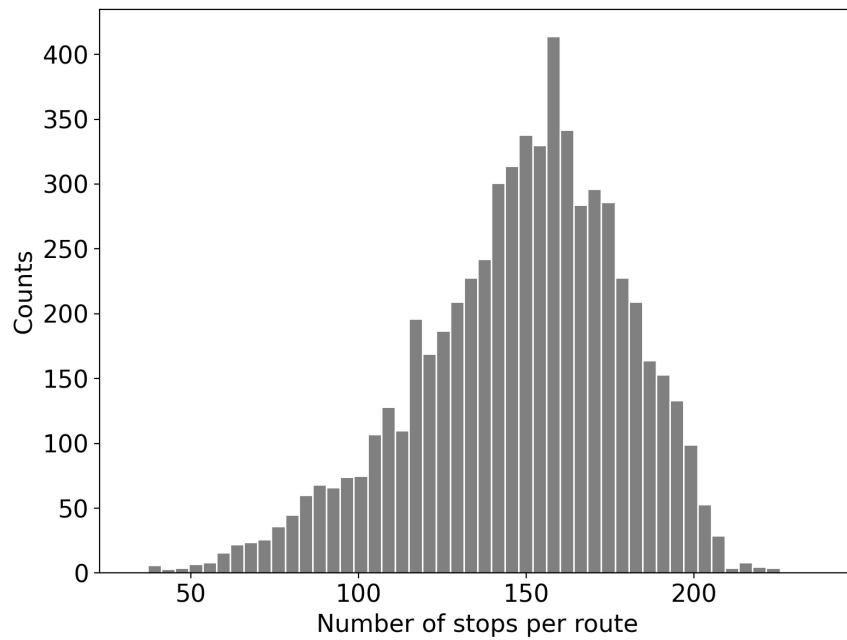


Figure 7: Relationship between stop sequence and zone sequence.

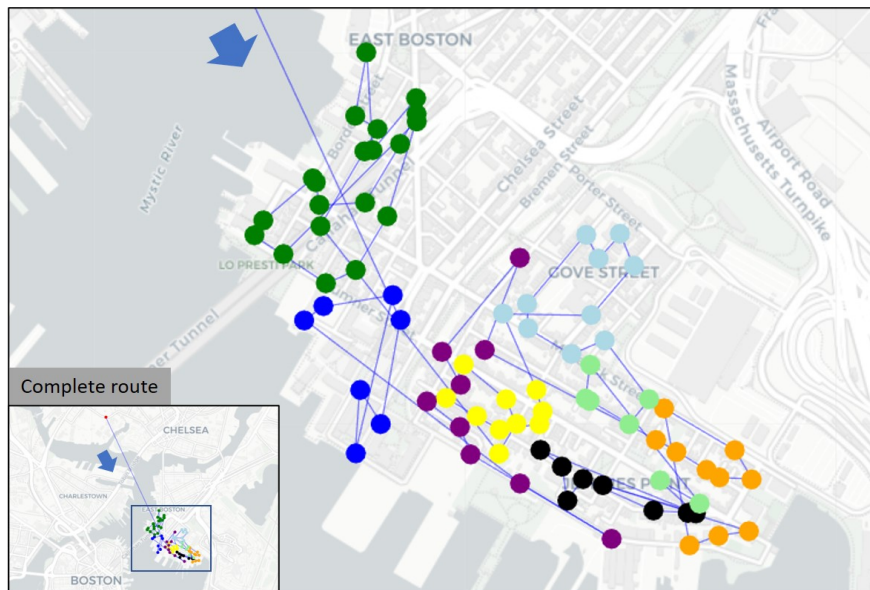
5.2. Experimental setup

We randomly select 4,889 routes for model training and cross-validation, and the remaining 1,223 routes are used to evaluate/test model performance.

We consider a one-layer LSTM for both the encoder and decoder with the hidden unit sizes of 32 (i.e., $K_h^D = K_e = K_h^E = K_d = 32$). And the ASNN is set with 2 hidden layers with 128 hidden units in each layer. We train the model using Adam optimizer with a default learning rate of 0.001 and 30 training epochs. To utilize the planned route information, the input zone sequence for the LSTM encoder is set as the TSP



(a) Number of stops distribution



(b) Actual route example

Figure 6: Description of dataset

result (i.e., lowest travel time). That is, the input sequence $(s_1, \dots, s_n) = (s_1^T, \dots, s_n^T)$.

In the case study, x_i represents zone features, including the latitude and longitude of the zone center, number of stops in the zone, number of intersections in the zone, number of packages in the zone, total service time in the zone, total package size in the zone, and the travel time from this zone to all other zones. The zone pair features $z_{(i)}^j$ includes the travel time from $\hat{s}_{(i)}$ to s_j and zone ID relationship characteristics. For example, the zone IDs “B-6.2C” and “B-6.3A” signal that they belong to the higher-level cluster “B-6”. As we assume all pair-wise features are captured by $z_{(i)}^j$, $\phi(x_{(i)}, x_j)$ is not specified in this case study.

Consistent with the Amazon Last Mile Routing Research Challenge, we evaluate the quality of the predicted stop sequences using a “disparity score” defined as follows:

$$R(A, B) = \frac{SD(A, B) \cdot ERP_{\text{norm}}(A, B)}{ERP_e(A, B)} \quad (12)$$

where $R(A, B)$ is the disparity score for the actual sequence A and predicted sequence B , and $SD(A, B)$ is the sequence deviation defined as

$$SD(A, B) = \frac{2}{n(n-1)} \sum_{i=2}^n (|c_{[B_i]} - c_{[B_{i-1}]}| - 1) \quad (13)$$

where n is the total number of stops, B_i is the i -th stop of sequence B , $c_{[B_i]}$ is the index of stop B_i in the actual sequence A (i.e., its position in sequence A). In the case of $A = B$ (i.e., perfectly predicted), we have $c_{[B_i]} - c_{[B_{i-1}]} = 1$ for all $i = 2, \dots, n$, and $SD(A, B) = 0$.

$ERP_{\text{norm}}(A, B)$ is the Edited Distance with Real Penalty (ERP) defined by the following recursive formula:

$$ERP_{\text{norm}}(A, B) = ERP_{\text{norm}}(A_{2:|A|}, B_{2:|B|}) + \text{TIME}_{\text{norm}}(A_1, B_1) \quad (14)$$

where $\text{TIME}_{\text{norm}}(s_i, s_j) = \frac{\text{TIME}(s_i, s_j)}{\sum_{j' \in \{1, \dots, n\}} \text{TIME}(s_i, s_{j'})}$ is the normalized travel time from stop s_i to stop s_j . $ERP_e(A, B)$ is the number of edit operations (insertions, substitutions, or deletions) required to transform sequence A to sequence B as when executing the recursive ERP_{norm} formulation. Hence, the ratio $\frac{ERP_{\text{norm}}(A, B)}{ERP_e(A, B)}$ represents the average normalized travel time between the two stops involved in each ERP edit operation. In the case of $A = B$, we have $\frac{ERP_{\text{norm}}(A, B)}{ERP_e(A, B)} = 0$.

The disparity score $R(A, B)$ describes how well the model-estimated sequence matches the known actual sequence. Lower score indicates better model performance. A score of zero means perfect prediction. The final model performance is evaluated by the mean score over all routes in the test set.

In addition to the disparity score, we also evaluate the prediction accuracy of the first four zones in each route. We choose the first four because the minimum number of zones in a route is four.

5.3. Benchmark models

The following optimization and machine learning models are used as benchmarks to compare with the proposed approach.

Conventional TSP. The first benchmark model is the zone sequence generated by conventional TSP, which we treat as the planned route with the lowest travel time.

ASNN model. The ASNN component can be trained to predict the next zone given the current zone, and the prediction sequence can be constructed in a greedy way starting from the given depot. The training zone

pairs (including from depot to the first zone) are extracted from all sequences in the training routes. And the input features are the same as the ASNN component in the proposed model except for $(d_{(i)}, e_j)$ (i.e., output vectors from LSTM decoder and encoder, respectively). All hyper-parameters of the ASNN model are the same as the attention component.

Inspired by the importance of the first zone, we also implement another sequence generation method similar to Section 4.3. That is, we go through all zones in a route and assume it is the first zone, then use the trained ASNN to predict the remaining sequence. The final sequence is selected as the one with the lowest travel time.

LSTM-encoder-decoder. The LSTM-encoder-decoder (LSTM-E-D) architecture is a typical seq2seq model proposed by [Sutskever et al. \(2014\)](#). The model structure is shown in Figure 8. In the decoder stage, the model outputs the predicted zone based on last predicted zone's information. The model formulation can be written as

$$h_i^E, e_i = \text{LSTM}(x_i, h_{i-1}^E; \theta^E) \quad \forall i = 1, \dots, n \quad (15)$$

$$h_{(i+1)}^D, d_{(i)} = \text{LSTM}(x_{(i)}, h_{(i)}^D; \theta^D) \quad \forall i = 0, 1, \dots, n \quad (16)$$

The decoder output vector $d_{(i)}$ are, then feed into a fully-connected (FC) layer to calculate probability of the next stop:

$$g_{(i)} = \text{FC}(d_{(i)}; \theta^{\text{FC}}) \quad \forall i = 1, \dots, n \quad (17)$$

$$\mathbb{P}(c_{i+1} \mid c_1, \dots, c_i, \mathcal{S}, X^{\mathcal{S}}; \theta) = \text{Softmax}(g_{(i)}) \quad \forall i = 1, \dots, n \quad (18)$$

where $g_{(i)} \in \mathbb{R}^{K_z}$, K_z is the maximum number of zones in the dataset. And the next predicted stop is selected by maximizing $\mathbb{P}(c_{i+1} = j \mid c_1, \dots, c_i, \mathcal{S}, X^{\mathcal{S}}; \theta)$ for all $s_j \in \mathcal{S} \setminus \mathcal{S}_{(i)}^V$ (i.e., the zones that are not in the route and that have been visited are excluded).

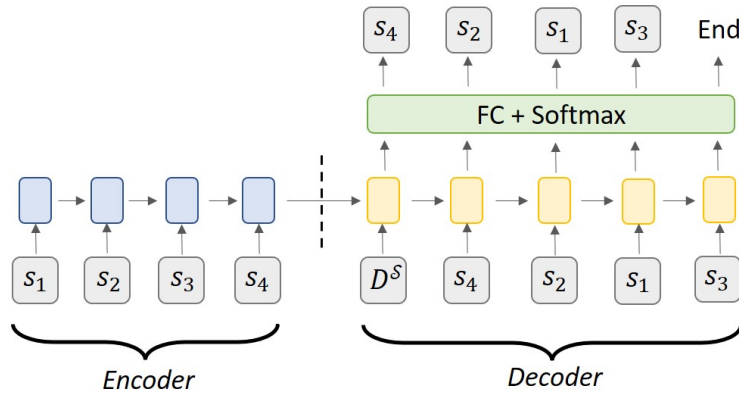


Figure 8: Model architecture of the LSTM-E-D seq2seq prediction model.

Original Pointer Network. Another benchmark model is the original pointer network (Pnt Net) proposed by [\(Vinyals et al., 2015\)](#). The overall architecture of the pointer network is similar to the proposed model

except for the attention component. Specifically, the pointer network calculates attention as:

$$u_{(i)}^j = W_1^T \tanh(W_2 e_j + W_3 d_{(i)}) \quad \forall i, j = 1, \dots, n \quad (19)$$

$$a_{(i)}^j = \frac{\exp(u_{(i)}^j)}{\sum_{j'=1}^n \exp(u_{(i)}^{j'})} \quad \forall i, j = 1, \dots, n \quad (20)$$

The original pointer network does not include the pair-wise local information $(z_{(i)}^j, \phi(x_{(i)}, x_j))$, and the attention calculation is only quantified from three learnable parameters W_1, W_2 , and W_3 , which may limit its capacity in prediction. We observe that the original pointer network without local information performs extremely badly. For a fair comparison, we add the local information with the similar format in Eq. 19 as:

$$u_{(i)}^j = W_1^T \tanh(W_2 e_j + W_3 d_{(i)}) + W_4 \begin{bmatrix} z_{(i)}^j \\ \phi(x_{(i)}, x_j) \end{bmatrix} \quad \forall i, j = 1, \dots, n \quad (21)$$

After training the model, we generate the final sequence with the greedy algorithm and Algorithm 1, respectively.

5.4. Results

5.4.1. Model comparison.

Table 1 presents the performance of different models. Note that for all approaches except for the TSP, we generate sequences based on two different methods (greedy and Algorithm 1) for comparison. The standard deviation of disparity scores is taken over all testing routes. Results show that sequence generation with Algorithm 1 (i.e., iterating different first zones) can consistently reduce the disparity score for all machine learning methods. It implies that the first zone prediction and the global view (i.e., shortest path) are important for estimating the driver's trajectory.

The proposed method outperforms all other models, both in disparity scores and prediction accuracy. This means the proposed pair-wise ASNN-based attention (Eq. 6) has better performance than the original content-based attention (Eq. 21). The comparison between LSTM-E-D and Pnt Net models demonstrates the effectiveness of the attention mechanism. All machine learning models except for LSTM-E-D can outperform the baseline TSP sequence with Algorithm 1 sequence generation method, suggesting that the hidden trajectory patterns can be learned from the data.

Another observation is that, the prediction accuracy and disparity score do not always move in the same direction. For example, the LSTM-E-D model with Algorithm 1 sequence generation, though has lower accuracy, shows a better disparity score. This is because the accuracy metric does not differentiate “how wrong an erroneous prediction is”. By the definition of disparity score, if a stop is s_i but the prediction is s_j , and s_j and s_i are geographically close to each other, the score does not worsen too much. This suggests a future research direction in using disparity score as the loss function (e.g., training by RL) instead of cross-entropy loss.

Figure 9 shows the distribution of disparity scores for our proposed method with Algorithm 1 sequence generation (i.e., the best model). We observe that the prediction performance varies a lot across different routes. There is a huge proportion of routes with very small disparity scores (less than 0.01). The mean score is impacted by outlier routes. The median score is 0.0340, which is smaller than the mean value.

Table 1: Model performance

Sequence generation	Model	Disparity score		Prediction accuracy			
		Mean	Std. Dev	1st zone	2nd zone	3rd zone	4th zone
-	TSP	0.0443	0.0289	0.207	0.185	0.163	0.168
Greedy	ASNN	0.0470	0.0289	0.150	0.141	0.119	0.123
	LSTM-E-D	0.0503	0.0313	0.207	0.183	0.161	0.166
	Pnt Net	0.0460	0.0309	0.224	0.204	0.186	0.165
	Ours	0.0417	0.0306	0.241	0.231	0.224	0.221
Algorithm 1	ASNN	0.0429	0.0299	0.221	0.213	0.203	0.195
	LSTM-E-D	0.0501	0.0305	0.182	0.156	0.142	0.149
	Pnt Net	0.0382	0.0301	0.286	0.273	0.262	0.274
	Ours	0.0369	0.0301	0.320	0.310	0.303	0.314

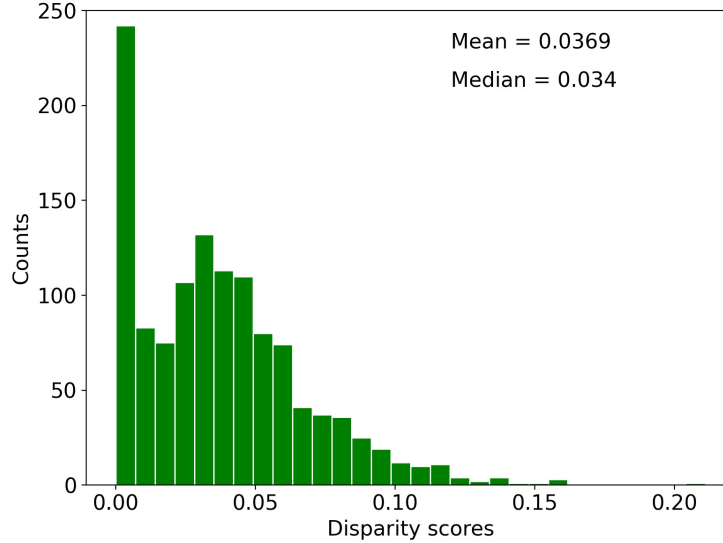


Figure 9: Disparity score distribution of the best model

5.4.2. Factors on trajectory predictability.

As our proposed model exhibits various levels of predictability across different routes, we aim to investigate which attributes of a route cause high (or low) predictability. This can be done by running a regression model with the disparity score as the dependent variable and route attributes (e.g., locations, departure time, package numbers) as independent variables. The variables used are defined as follows:

- Total planned service time: The estimated time to deliver all packages in the route (service time only, excluding travel time).
- Earliest time window constraint: The earliest due time to deliver packages with time window constraint minus the vehicle departure time. The smaller the value, the tighter the time limit.
- Avg. # traffic signals: Average number of traffic signals in each zone of the route (obtained from OpenStreetMap data).

- If high-quality route: A dummy variable indicating whether the route is labeled as “high quality” by Amazon or not (Yes = 1). High quality means the actual travel time of the route is similar to or better than Amazon’s expectation.
- If in LOCATION: A dummy variable indicating whether the route is in a specific city or not (Yes = 1).
- If departure TIME: A dummy variable indicating the (local) departure time (e.g., before 7AM, after 10AM).

Table 2 shows the results of the regression. Since the dependent variable is *disparity scores*, a negative sign indicates a positive impact on the predictability. We observe that routes with tighter time window constraints and more stops are easier to predict. This may be due to the fact that these routes are usually harder to deliver. Hence, to avoid the risk of violating time constraints or delay, drivers tend to follow the planned routes and thus the route sequences are easier predict. We also find that routes associated with larger vans (i.e., larger vehicle capacity) are more predictable. The reason may be that larger vans are less flexible in choosing different routes, thus drivers are more likely to follow the navigation. Another important factor for better predictability is high-quality routes. This may be because high-quality routes are closer to the TSP sequence which we use as inputs. Finally, routes in LA are more predictable than in other areas such as Chicago and Boston.

Table 2: Factors on trajectory predictability

Variables	Coefficients ($\times 10^{-3}$)	Variables	Coefficients ($\times 10^{-3}$)
Intercept	91.07 **	If high quality route	-1.66×10^{-14} **
Total # of packages	0.059	If in LA	-4.998 *
Total planned service time	-0.476	If in Chicago	0.783
Earliest time window constraint	-3.047 **	If in Boston	-3.354
Avg. # traffic signals	-3.255	If on weekends	1.775
Total # of stops	-0.142 **	If departure before 7AM	0.582
Vehicle capacity (m^3)	-6.041 *	If departure after 10AM	-2.704

Number of routes: 1,002.

R^2 : 0.065;

** : p -value < 0.01 ; * : p -value < 0.05 .

5.4.3. Impact of input sequence.

All machine learning models in Table 1 (except for ASNN) have the LSTM encoder component, which requires the specification of input zone sequence. As mentioned in Section 5.2, we currently use the TSP sequence as input. It is worth exploring the model performance if we use a random zone sequence instead, which corresponds to the scenario without planned route information. Table 3 shows the model performance without the TSP sequence information. Since the ASNN result does not rely on TSP information, it is not listed in the table. Results show that the LSTM-E-D model becomes much worse with a random sequence as inputs, while the performance of Pnt Net and our method is only slightly affected. Even without the planned route information, the proposed model can still provide a reasonable estimation of driver trajectories.

Table 3: Model performance without TSP information

Sequence generation	Model	Disparity score		Prediction accuracy			
		Mean	Std. Dev	1st zone	2nd zone	3rd zone	4th zone
Greedy	LSTM-E-D	0.1176	0.0498	0.045	0.047	0.041	0.050
	Pnt Net	0.0512	0.0323	0.090	0.096	0.097	0.096
	Ours	0.0426	0.0311	0.204	0.192	0.195	0.196
Algorithm 1	LSTM-E-D	0.1054	0.0463	0.103	0.061	0.049	0.052
	Pnt Net	0.0398	0.0311	0.298	0.284	0.273	0.273
	Ours	0.0376	0.0307	0.316	0.298	0.302	0.298

5.5. Summary

Our numerical results show that our proposed model outperforms its benchmarks in terms of disparity scores and prediction accuracy, meaning that it can better predict the actual route trajectories taken by drivers. The comparison with benchmark models shows that our proposed ASNN-based pair-wise attention mechanism and our sequence generation algorithm (Algorithm 1) are both helpful for the prediction. Moreover, we can observe that the predictive performance varies across different routes. Factors such as route quality, delivery time windows, and the total number of stops of a route affect predictability. Finally, the proposed model is insensitive to the input sequence. The prediction performance only slightly decreases when the input sequence is changed from the TSP solution to a random stop sequence. This property implies that we only need the set of stops to implement the model and obtain high-quality solution, while information on the planned route sequence is not strictly required.

6. Conclusion and Future Research

In this paper, we propose a pair-wise attention-based pointer neural network that predicts actual driver trajectories on last-mile delivery routes for given sets of delivery stops. Compared to previously proposed pointer networks, this study leverages a new alternative specific neural network-based attention mechanism to incorporate pair-wise local information (such as relative distances and locations of stops) for the attention calculation. To better capture the global efficiency of a route in terms of operational cost (i.e., total travel time), we further propose a new sequence generation algorithm that finds the lowest-cost route sequence by iterating through different first stops.

We apply our proposed method to a large set of real operational route data provided by the Amazon Last-Mile Routing Research Challenge in 2021. The results show that our proposed method can outperform a wide range of benchmark models in terms of both the disparity score and prediction accuracy, meaning that the predicted route sequence is closer to the actual sequence executed by drivers. Compared to the best benchmark model (original pointer network), our method reduces the disparity score from 0.0382 to 0.0369, and increases the average prediction accuracy of the first four zones from 0.229 to 0.312. Moreover, our proposed sequence generation method can consistently improve the prediction performance for all models. The disparity scores are reduced by 10-20% across different models. Lastly, we show that the proposed methodology is robust against changes in the input sequence pattern. Compared to an optimal TSP solution as the input sequence, a random input sequence only slightly increases the disparity score from 0.0369 to 0.0376.

The data-driven route planning method proposed in this paper has several highly relevant practical implications. First, our proposed model performs well at predicting stop sequences that would be preferable to delivery drivers in a real operational environment, even if it is not provided with a theoretically optimal (i.e., minimal route duration) planned TSP sequence as an input. Therefore, the model can be used to generate a predicted actual stop sequence that a driver would likely be taking for a given set of delivery stops. The prediction can serve as a new ‘empirical’ planned route that is informed by historical driver behavior and thus more consistent with the driver’s experience and preferences. Second, by comparing the stop sequence predicted by our model with the traditional, TSP-based planned stop sequence, a route planner may infer potential reasons for the drivers’ deviations and adjust the company’s planning procedures and/or driver incentives if necessary. Third, as stop sequence generation using machine learning models is computationally more efficient than traditional optimization-based approaches, a trained machine learning model can be applied in real-time to quickly re-optimize routes when drivers are unexpectedly forced to deviate from their original stop sequence (e.g., due to road closures) and need updated routing strategies.

Based on the work presented in this paper, a number of fruitful future research avenues arise. First, instead of focusing on stop sequence prediction, future work may improve the interpretability of such prediction models and develop machine learning approaches that better explain which factors cause drivers to deviate from a planned stop sequence and how they affect their actual route trajectories. Second, future work should attempt to combine the strengths of optimization-based route planning approaches and machine learning by incorporating tacit driver knowledge learned via machine learning models into route optimization algorithms.

Appendices

Appendix A. Mathematical Formulation of a LSTM Cell

The details of an LSTM cell, $h_t, e_t = \text{LSTM}(x_t, h_{t-1}; \theta)$, is shown below:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (\text{A.1})$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (\text{A.2})$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (\text{A.3})$$

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (\text{A.4})$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (\text{A.5})$$

$$h_t = o_t \circ \sigma_h(c_t) \quad (\text{A.6})$$

$$e_t = h_t \text{ (if this is a single layer one-directional LSTM)} \quad (\text{A.7})$$

where $[W_f, W_i, W_o, W_c, U_f, U_i, U_o, U_c, b_f, b_i, b_o, b_c] = \theta$ is the vector of learnable parameters. x_t is the input vector to the LSTM unit. f_t is the forget gate’s activation vector. i_t is the input/update gate’s activation vector. o_t is the output gate’s activation vector. h_t is the hidden state vector. e_t is the output vector of the LSTM. Note that for a multi-layer or bidirectional LSTM, e_t may not equal to h_t . In this study, we use a single layer one-directional LSTM and thus have $e_t = h_t$. More details on the output vector can be found in [Pytorch \(2021\)](#). \tilde{c}_t is the cell input activation vector. c_t is the cell state vector. “ \circ ” indicates the component-wise multiplication.

Appendix B. From Zone Sequence to Stop Sequence

The complete stop sequence is generated based on the given zone sequence. The detailed generation process is shown in Algorithm 2.

Algorithm 2 Complete sequence generation. Input: zone sequence $(\hat{z}_{(1)}, \dots, \hat{z}_{(n)})$, depot D^S , set of stops in each zone $\mathcal{S}_{(i)}, i = 1, \dots, n$. $\text{PATHTSP}(\mathcal{S}, s_{first}, s_{last})$ and $\text{TOURTSP}(\mathcal{S})$ are two oracle functions for solving path and tour TSP problems given the set of stops \mathcal{S} , first stop s_{first} and last stop s_{last} to be visited.

```

1: function COMPLETESEQGENERATION( $(\hat{z}_{(1)}, \dots, \hat{z}_{(n)}), \{\mathcal{S}_{(i)}, i = 1, \dots, n\}$ )
2:    $s_{prev} \leftarrow D^S$ 
3:    $s_{complete}^* \leftarrow (s_{prev})$  ▷ Initialize the complete stop sequence with depot
4:   for  $i \in \{1, \dots, n-1\}$  do
5:      $\mathcal{S}_{first} \leftarrow$  Set of three stops in  $\mathcal{S}_{(i)}$  that are closest to  $s_{prev}$ 
6:      $\mathcal{S}_{last} \leftarrow$  Set of three stops in  $\mathcal{S}_{(i)}$  that are closest to all stops in  $\mathcal{S}_{i+1}$  on average
7:      $\mathcal{P}_{(i)} \leftarrow \emptyset$  ▷ Initialize the set of optimal paths in zone  $\hat{z}_{(i)}$ 
8:     for  $s_{first} \in \mathcal{S}_{first}$  do
9:       for  $s_{last} \in \mathcal{S}_{last}$  do
10:        if  $s_{first} = s_{last}$  then
11:           $\hat{p}_{temp}, t_{temp} = \text{TOURTSP}(\mathcal{S}_{(i)})$  ▷ Solve the optimal tour and travel time for zone  $\hat{z}_{(i)}$ 
12:          Delete the last edge back to  $s_{first}$  in the tour  $\hat{p}_{temp}$ . Let the new path and travel time be  $\hat{p}'_{temp}$ 
            and  $t'_{temp}$ 
13:          Add  $\hat{p}'_{temp}$  and  $t'_{temp}$  to  $\mathcal{P}_{(i)}$ 
14:        else
15:           $\hat{p}_{temp}, t_{temp} = \text{PATHTSP}(\mathcal{S}_{(i)}, s_{first}, s_{last})$  ▷ Solve the optimal path and travel time for zone  $i$ 
16:          Add  $\hat{p}_{temp}$  and  $t_{temp}$  to  $\mathcal{P}_{(i)}$ 
17:         $\hat{p}_{(i)} \leftarrow$  Path in  $\mathcal{P}_{(i)}$  with the minimum travel time
18:         $s_{complete}^* \leftarrow (s_{complete}^*, \hat{p}_{(i)})$  ▷ Concatenate two sequence
19:         $s_{prev} \leftarrow$  Last stop of path  $\hat{p}_{(i)}$ 
20:         $s_{complete}^* \leftarrow (s_{complete}^*, D^S)$  ▷ Concatenate the last stop as the depot
21:   return  $s_{complete}^*$ 

```

Consider an optimal zone sequence, $(\hat{z}_{(1)}, \dots, \hat{z}_{(n)})$, generated from the proposed machine learning method. We can always add the depot before the first and after last zone (i.e., $(D^S, \hat{z}_{(1)}, \dots, \hat{z}_{(n)}, D^S)$) and make the whole zone sequence a loop. For each zone $\hat{z}_{(i)}$, we aim to generate a within-zone path $\hat{p}_{(i)}$, and the final stop sequence will be $(D^S, \hat{p}_{(1)}, \dots, \hat{p}_{(n)}, D^S)$.

When generating $\hat{p}_{(i)}$ for zone $\hat{z}_{(i)}$, we assume $\hat{p}_{(i-1)}$ is known (generated from the last step and $\hat{p}_{(0)} = (D^S)$). Let the set of all stops in zone $\hat{z}_{(i)}$ be $\mathcal{S}_{(i)}$. We identify three *potential* first stops and last stops of path $\hat{p}_{(i)}$ based on following rules:

- Three potential first stops of $\hat{p}_{(i)}$ are the three most closest stops (in travel time) to $\hat{p}_{(i-1)}$'s last stop.
- Three potential last stops of $\hat{p}_{(i)}$ are the three most closest stops (in travel time) to all stops in $\mathcal{S}_{(i+1)}$ on average. Note that $\mathcal{S}_{(n+1)} = \{D^S\}$

With three potential first stops and last stops, we then solve path TSP problems between any first and last stop pair to generate the potential optimal inner zone path with the shortest travel time. In this step, at most

nine small-scale path TSP problems will be solved since there might be overlapping between the first and the last stops. If the first and the last stops are identical, we solve a tour TSP problem and output the path by deleting the last edge which traverses back to the first stop in the tour.

After having all potential inner zone paths and total path travel time between any first and last stop pair, we keep the path with the minimum travel time as the inner zone sequence, $\hat{p}_{(i)}$. The key assumption we make here about drivers is that they will deliver packages within a zone following a path that minimizes their total travel time. With the optimal inner zone stop sequence of the current zone, we then move to the next visited zone in the optimal zone sequence and repeat the same procedure until we generate the complete stop sequence.

References

- Applegate, D.L., Bixby, R.E., Chvatal, V., Cook, W.J., 2006. The Travelling Salesman Problem. Princeton University Press. URL: <http://www.jstor.org/stable/j.ctt7s8xg.9>.
- Bahdanau, D., Cho, K.H., Bengio, Y., 2015. Neural machine translation by jointly learning to align and translate, in: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.
- Bello, I., Pham, H., Le, Q.V., Norouzi, M., Bengio, S., 2019. Neural combinatorial optimization with reinforcement learning. 5th International Conference on Learning Representations, ICLR 2017 - Workshop Track Proceedings , 1–15.
- Cheikhrouhou, O., Khoufi, I., 2021. A comprehensive survey on the multiple traveling salesman problem: Applications, approaches and taxonomy. Computer Science Review 40, 100369. URL: <https://doi.org/10.1016/j.cosrev.2021.100369>, doi:10.1016/j.cosrev.2021.100369.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation, in: Conference on Empirical Methods in Natural Language Processing (EMNLP 2014).
- Chorowski, J., Bahdanau, D., Serdyuk, D., Cho, K., Bengio, Y., 2015. Attention-based models for speech recognition. arXiv preprint arXiv:1506.07503 .
- Chung, J., Gulcehre, C., Cho, K., Bengio, Y., 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555 .
- Cortes, J.D., Suzuki, Y., 2021. Last-mile delivery efficiency: en route transloading in the parcel delivery industry. International Journal of Production Research 0, 1–18. URL: <https://doi.org/00207543.2021.1907628>, doi:10.1080/00207543.2021.1907628.
- Davendra, D., Bialic-Davendra, M., 2020. Introductory chapter: Traveling salesman problem - an overview, in: Novel Trends in the Traveling Salesman Problem. IntechOpen. URL: <https://doi.org/10.5772/intechopen.94435>, doi:10.5772/intechopen.94435.
- Graves, A., 2013. Generating Sequences With Recurrent Neural Networks URL: <http://arxiv.org/abs/1308.0850>.
- Halim, A.H., Ismail, I., 2017. Combinatorial optimization: Comparison of heuristic algorithms in travelling salesman problem. Archives of Computational Methods in Engineering 26, 367–380. URL: <https://doi.org/10.1007/s11831-017-9247-y>, doi:10.1007/s11831-017-9247-y.
- Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation 9, 1735–1780. URL: <https://doi.org/10.1162/neco.1997.9.8.1735>, doi:10.1162/neco.1997.9.8.1735, arXiv:<https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf>.

- Huang, C.Z.A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., Dai, A.M., Hoffman, M.D., Dinculescu, M., Eck, D., 2018. Music transformer. arXiv preprint arXiv:1809.04281 .
- Jaller, M., Holguín-Veras, J., Hodge, S., 2013. Parking in the city. *Transportation Research Record* , 46–56doi:[10.3141/2379-06](https://doi.org/10.3141/2379-06).
- Jonker, R., Volgenant, T., 1983. Transforming asymmetric into symmetric traveling salesman problems. *Operations Research Letters* 2, 161–163. URL: [https://doi.org/10.1016/0167-6377\(83\)90048-2](https://doi.org/10.1016/0167-6377(83)90048-2), doi:[10.1016/0167-6377\(83\)90048-2](https://doi.org/10.1016/0167-6377(83)90048-2).
- Joshi, C.K., Laurent, T., Bresson, X., 2019. On Learning Paradigms for the Travelling Salesman Problem. *Advances in Neural Information Processing Systems* , 1–9URL: <http://arxiv.org/abs/1910.07210>.
- Karatzoglou, A., Jablonski, A., Beigl, M., 2018. A seq2seq learning approach for modeling semantic trajectories and predicting the next location, in: *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Association for Computing Machinery, New York, NY, USA. p. 528–531. URL: <https://doi.org/10.1145/3274895.3274983>, doi:[10.1145/3274895.3274983](https://doi.org/10.1145/3274895.3274983).
- Kool, W., Van Hoof, H., Welling, M., 2019. Attention, learn to solve routing problems! *7th International Conference on Learning Representations, ICLR 2019* , 1–25.
- Küçükoğlu, İ., Dewil, R., Cattryse, D., 2019. Hybrid simulated annealing and tabu search method for the electric travelling salesman problem with time windows and mixed charging rates. *Expert Systems with Applications* 134, 279–303. URL: <https://doi.org/10.1016/j.eswa.2019.05.037>, doi:[10.1016/j.eswa.2019.05.037](https://doi.org/10.1016/j.eswa.2019.05.037).
- Liang, Z., Du, J., Li, C., 2020. Abstractive social media text summarization using selective reinforced seq2seq attention model. *Neurocomputing* 410, 432–440.
- Lim, S.F.W., Winkenbach, M., 2019. Configuring the last-mile in business-to-consumer e-retailing. *California Management Review* 61, 132–154.
- Liu, S., Jiang, H., Chen, S., Ye, J., He, R., Sun, Z., 2020. Integrating dijkstra’s algorithm into deep inverse reinforcement learning for food delivery route planning. *Transportation Research Part E: Logistics and Transportation Review* 142, 102070. URL: <https://www.sciencedirect.com/science/article/pii/S1366554520307213>, doi:<https://doi.org/10.1016/j.tre.2020.102070>.
- Liu, T., Wang, K., Sha, L., Chang, B., Sui, Z., 2018. Table-to-text generation by structure-aware seq2seq learning, in: *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Lu, Y., Rai, H., Chang, J., Knyazev, B., Yu, G., Shekhar, S., Taylor, G.W., Volkovs, M., 2021. Context-aware scene graph generation with seq2seq transformers, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 15931–15941.
- Luong, M.T., Pham, H., Manning, C.D., 2015. Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025 .
- Ma, Q., Ge, S., He, D., Thaker, D., Drori, I., 2019. Combinatorial Optimization by Graph Pointer Networks and Hierarchical Reinforcement Learning URL: <http://arxiv.org/abs/1911.04936>.
- Matai, R., Singh, S., Lal, M., 2010. Traveling salesman problem: an overview of applications, formulations, and solution approaches, in: *Traveling Salesman Problem, Theory and Applications*. InTech. URL: <https://doi.org/10.5772/12909>, doi:[10.5772/12909](https://doi.org/10.5772/12909).
- McKinsey & Company, 2021. How e-commerce share of retail soared across the globe: A look at eight countries. URL: <https://www.mckinsey.com/featured-insights/coronavirus-leading-through-the-crisis/charting-the-path-to-the-next-normal/>

- [how-e-commerce-share-of-retail-soared-across-the-globe-a-look-at-eight-countries](#).
- Merchán, D., Arora, J., Pachon, J., Konduri, K., Winkenbach, M., Parks, S., Noszek, J., 2022. 2021 amazon last mile routing research challenge: Data set. *Transportation Science* .
- Mladenović, N., Todosijević, R., Urošević, D., 2012. An efficient GVNS for solving traveling salesman problem with time windows. *Electronic Notes in Discrete Mathematics* 39, 83–90. URL: <https://doi.org/10.1016/j.endm.2012.10.012>, doi:10.1016/j.endm.2012.10.012.
- Pitney Bowes, 2020. Pitney Bowes Parcel Shipping Index. URL: <https://www.pitneybowes.com/us/shipping-index.html>.
- postnord, 2021. E-commerce in Europe 2020 - How the pandemic is changing e-commerce in Europe. Technical Report. URL: <https://www.postnord.se/siteassets/pdf/rapporter/e-commerce-in-europe-2020.pdf>.
- Purkayastha, R., Chakraborty, T., Saha, A., Mukhopadhyay, D., 2020. Study and analysis of various heuristic algorithms for solving travelling salesman problem—a survey, in: *Advances in Intelligent Systems and Computing*. Springer Singapore, pp. 61–70. URL: https://doi.org/10.1007/978-981-15-2188-1_5, doi:10.1007/978-981-15-2188-1_5.
- Pytorch, 2021. Pytorch LSTM document. URL: <https://pytorch.org/docs/stable/generated/torch.nn.LSTM.html>.
- Ren, S., Choi, T.M., Lee, K.M., Lin, L., 2020. Intelligent service capacity allocation for cross-border e-commerce related third-party-forwarding logistics operations: A deep learning approach. *Transportation Research Part E: Logistics and Transportation Review* 134, 101834. URL: <https://www.sciencedirect.com/science/article/pii/S1366554519311688>, doi:<https://doi.org/10.1016/j.tre.2019.101834>.
- Rose, W.J., Mollenkopf, D.A., Autry, C., Bell, J., 2016. Exploring urban institutional pressures on logistics service providers. *International Journal of Physical Distribution & Logistics Management* 46. doi:10.1108/09600035199500001.
- Salman, R., Ekstedt, F., Damaschke, P., 2020. Branch-and-bound for the precedence constrained generalized traveling salesman problem. *Operations Research Letters* 48, 163–166.
- da Silva, R.F., Urrutia, S., 2010. A general VNS heuristic for the traveling salesman problem with time windows. *Discrete Optimization* 7, 203–211. URL: <https://doi.org/10.1016/j.disopt.2010.04.002>, doi:10.1016/j.disopt.2010.04.002.
- Snoeck, A., Winkenbach, M., 2021. A discrete simulation-based optimization algorithm for the design of highly responsive last-mile distribution networks. *Transportation Science* .
- Sutskever, I., Vinyals, O., Le, Q.V., 2014. Sequence to sequence learning with neural networks, in: *Advances in Neural Information Processing Systems*, pp. 3104–3112.
- Traub, V., Vygen, J., Zenklusen, R., 2021. Reducing path TSP to TSP. *SIAM Journal on Computing* , STOC20–24–STOC20–53URL: <https://doi.org/10.1137/20m135594x>, doi:10.1137/20m135594x.
- United Nations Department of Economic and Social Affairs, 2019. World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420). Technical Report. New York: United Nations. URL: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>.
- US Census Bureau, 2021. Quarterly e-commerce retail sales 2nd quarter 2021. Technical Report. U.S. Census Bureau of the Department of Commerce. URL: <http://www2.census.gov/retail/releases/historical/ecommm/07q4.pdf>.

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need. *Advances in Neural Information Processing Systems 2017-Decem*, 5999–6009. URL: <http://arxiv.org/abs/1706.03762>, [arXiv:1706.03762](https://arxiv.org/abs/1706.03762).
- Vinyals, O., Bengio, S., Kudlur, M., 2016. Order matters: Sequence to sequence for sets. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 1–11.
- Vinyals, O., Meire, F., Navdeep, J., 2015. Pointer Networks. *Advances in Neural Information Processing Systems*, 1–9.
- Wang, S., Cao, J., Chen, H., Peng, H., Huang, Z., 2020a. Seqst-gan: Seq2seq generative adversarial nets for multi-step urban crowd flow prediction 6. URL: <https://doi.org/10.1145/3378889>, doi:10.1145/3378889.
- Wang, S., Mo, B., Zhao, J., 2020b. Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions. *Transportation Research Part C: Emerging Technologies* 112, 234–251. doi:10.1016/j.trc.2020.01.012.
- Winkenbach, M., Parks, S., Noszek, J., 2021. Technical Proceedings of the Amazon Last Mile Routing Research Challenge URL: <https://dspace.mit.edu/handle/1721.1/131235>.
- World Economic Forum, 2020. The Future of the Last-Mile Ecosystem. Technical Report January. URL: <https://www.weforum.org/reports/the-future-of-the-last-mile-ecosystem>.
- Wu, R., Zhuang, Y., Xu, K., Zhang, H., Chen, B., 2020. Pq-net: A generative part seq2seq network for 3d shapes, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Xu, Z., Wang, S., Zhu, F., Huang, J., 2017. Seq2seq fingerprint: An unsupervised deep molecular embedding for drug discovery, in: *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, Association for Computing Machinery, New York, NY, USA*. p. 285–294. URL: <https://doi.org/10.1145/3107411.3107424>, doi:10.1145/3107411.3107424.
- Yuan, Y., Cattaruzza, D., Ogier, M., Semet, F., 2020. A branch-and-cut algorithm for the generalized traveling salesman problem with time windows. *European Journal of Operational Research* 286, 849–866.
- Zhang, Y., Li, D., Wang, Y., Fang, Y., Xiao, W., 2019. Abstract text summarization with a convolutional seq2seq model. *Applied Sciences* 9. URL: <https://www.mdpi.com/2076-3417/9/8/1665>, doi:10.3390/app9081665.
- Zhang, Y., Li, Y., Zhang, G., 2020. Short-term wind power forecasting approach based on seq2seq model using nwp data. *Energy* 213, 118371. URL: <https://www.sciencedirect.com/science/article/pii/S036054422031478X>, doi:<https://doi.org/10.1016/j.energy.2020.118371>.