A review of methods to model route choice behavior of bicyclists: inverse reinforcement learning in spatial context and recursive logit

Thomas Koch
Centrum Wiskunde en Informatica
Amsterdam, The Netherlands
koch@cwi.nl

Elenna Dugundji
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands
e.r.dugundji@vu.nl

ABSTRACT

Used for route choice modeling by the transportation research community, recursive logit is a form of inverse reinforcement learning, the field of learning an agent's objective by observing it's behavior. By solving a large-scale system of linear equations it allows estimation of an optimal (negative) reward function in a computationally efficient way that performs for large networks and a large number of observations. In this paper we review examples of IRL models applied to real world travel trajectories and look at some of the challenges with recursive logit for modeling bicycle route choice in the city center area of Amsterdam.

CCS CONCEPTS

• **Computing methodologies → Inverse reinforcement learning**; • **Applied computing → Transportation**.

KEYWORDS

inverse reinforcement learning, GPS trajectory, route choice modeling, maximum entropy, dynamic discrete choice, markov decision process, dynamic programming, recursive logit, bicycle route behavior

ACM Reference Format:

Thomas Koch and Elenna Dugundji. 2020. A review of methods to model route choice behavior of bicyclists: inverse reinforcement learning in spatial context and recursive logit. In *3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation (GeoSim'20), November 3–6, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3423335.3428165

1 INTRODUCTION

Bicycling in Amsterdam is serious business, more than a third of journeys by Amsterdam residents are done on a bicycle. As the bicycle market share further increases and the number of Amsterdam residents keeps growing, problems such as traffic jams of bicycle near traffic lights start appearing. This makes it more and more important to model bicycle traffic for traffic studies and simulate potential policy changes. To do so it is important to understand the factor that drive bicycle route choice. In earlier work Koch et al. [10] we found that more factors than distance play a role bicyclists, meaning that other factors such as environment (through parks or next to water) or dedicated bicycle infrastructure might be able to explain route choice.

We start this paper with a thorough review of literature on using inverse reinforcement learning in spatial context and literature on recursive logit. We perform a case study to better understand bicyclist route choice behavior using a data-set of GPS traces collected by a large panel of volunteers. We conclude this paper with a reflection on challenges we encountered.

2 BACKGROUND

2.1 Inverse reinforcement learning on real world travel trajectories

Inverse reinforcement learning (IRL) aims to find reward function parameters θ by observing the behaviour of each agent in a Markov Decision Process (MDP) with a finite set S of N states. The reward function $R(\zeta)$ for trajectory $\zeta = \{s, a\}$, performing action a at state s with f_s the feature vector of state s, is given by:

$$R(\zeta) = \theta^T \mathbf{f}_\zeta = \sum_{s \in \zeta} \theta^T \mathbf{f}_s \tag{1}$$

.

In the computer science literature there are several studies performing IRL on real world problems. Ziebart et al. [20] introduced Maximum Entropy Inverse Reinforcement Learning in 2008 based on the principle of maximum entropy by Jaynes [9] that the probability of a trajectory ζ with higher reward is exponentially higher than that of a smaller reward: $P(\zeta) \propto e^{R(\zeta)}$. In order to learn from observed behaviour, the Maximum Entropy IRL algorithm maximizes the likelihood of the observed trajectories under the maximum entropy distribution T

$$\theta^* = \underset{\theta}{\operatorname{argmax}} \sum_{\zeta} \log P(\widetilde{\zeta} | \theta, T) \tag{2}$$

The maximum entropy distribution T is derived using

$$P(\zeta|\theta) = \frac{1}{Z(\theta)} e^{\sum s_j \in \zeta_i^{\theta \top \mathrm{f}_{s_j}}} \tag{3}$$

For parameters θ the partition function $Z(\theta)$ will always converge for the problem with finite horizons and infinite horizon problems with discounted reward weights. Since function 2 is convex for a deterministic MDP, gradients for optimizers can be obtained by taking the difference between the observed feature counts and the expected feature counts based on a given set of parameters θ, that

GeoSim'20, November 3–6, 2020, Seattle, WA, USA

ACM ISBN 978-1-4503-8161-1/20/11…$15.00
https://doi.org/10.1145/3423335.3428165

can be formulated as the expected state visitation frequencies D_{s_i}. To compute the gradients Ziebart et al. [20] uses:

$$\nabla L(\theta) = \widetilde{\mathrm{f}} - \sum_{\zeta} P(\zeta|\theta, T) f_{\zeta} = \widetilde{\mathrm{f}} - \sum_{s_i} D_{s_i} \mathrm{f}_{s_i} \tag{4}$$

To efficiently compute the expected state frequencies for parameters θ, Ziebart et al. [20] has proposed an algorithm that approximates the state frequencies by recursively backing up from each possible terminal state, computing each probability mass of each branch along the way, computing partition function Z at each action and state. The branching values give the local action probability that can be used to compute state frequencies and summed up for total frequency counts. Ziebart et al. [20] apply the Maximum Entropy IRL model to learn the reward function of taxi drivers on the road network of Pittsburgh, Pennsylvania. To do so, GPS logging of approximately 7403 trajectories are used to determine the cost of different road type, speed, number of lanes and turn costs. The MDP modeled from the road-network of Pittsburg is assumed to be deterministic with over 300,000 states (street segments) and 900,000 actions (transitions at intersections).

Hirakawa et al. [8] use maximum entropy IRL to learn from bird behaviour. As birds are equipped with GPS loggers but gaps may occur due to unavoidable issues with the equipment, a method is needed to fill those gaps with the most likely trajectory. By using maximum entropy IRL to find the reward function they determine the most likely route taken by bids based on environmental features. They applied this approach on one type of bird and found improvement over existing interpolation methods. The IRL model uses 53 trajectories in a 3 dimensional grid world, with 600 cells in height for 200 by 300 grid cells each a square of approximately 3 kilometers wide.

In Nguyen et al. [17] a generalization of the IRL problem is proposed that allows multiple locally consistent reward functions to generate the trajectories. By representing the IRL problem with a probabilistic graph model, an expectation-maximization (EM) algorithm can be devised to iteratively learn different reward functions and the stochastic transitions between them, in order to improve the likelihood of the observed trajectories. As a result, the EM algorithm can be used to derive locally consistent reward functions. Nguyen et al. [17] empirically evaluated their algorithm with a small real world network and GPS data of 59 taxis in Singapore. In this evaluation the road network is modelled as a simplified grid world with 193 states.

Mai et al. [12] proposes a generalized version of the causal entropy maximization problem, allowing the possibility to generate a class of maximum entropy IRL models. Their proposed generalized model has the advantage of being able to recover an expert function that would (partially) capture the impact of the connecting structure of the states on experts' decision. Their empirical evaluation on a real-world dataset and a grid-world dataset shows that their generalized model outperforms classical approaches in terms of recovering reward functions and demonstrated trajectories.

Mai et al. [14] proposes a tractable approach to compute directly a log-likelihood of observed trajectories with incomplete/missing data. By performing the training by solving a sequence of linear equations that does not depend on the number of missing segments it is efficient at handling a large number of missing segments. Their empirical evaluation showed that their approach outperforms other approaches.

Mo [16] looks at bicycle route choice applying the maximum entropy IRL approach. To achieve multi-reward functions an extension is used known as Behaviour Clustering IRL (BCIRL). He performs multiple experiments to investigate the applicability of these methods in the context of bicycle route choice. In this study it was found that a low number of demonstrated trajectories, short trajectory lengths, large number of Markov decision processes to be solved, and class imbalance were problematic issues for the methods. An application was performed on a dataset of GPS trajectories in Amsterdam, but no factors other than distance were found to be relevant.

In 2017 Wu et al. [18] proposed a data driven method that construct a MDP that models the decision making process of a public transit rider, decisions such as mode choice, route choice and transfer location choice. The purpose is to predict public transit route choice for urban planners, given various proposed transit construction scenarios. Using this MDP they use maximum entropy IRL to infer the passenger reward function from observed public transit chip card data (AFC) from Shenzhen, China for a period of 3 months. They model the real world as a grid world, dividing the world into grids of a square kilometer and the action set as the possible choice set of different bus and subway routes between each grid cell. The features they include are variables such as fare, travel time, number of transfers and the amount of time remaining to 9 am. In their study they find that they can find a reward function very closed to what is observed with regard to behaviour by public transit users and claim that it justifies their hypothesis that public transit users make sub optimal decisions.

In 2018 Wu et al. [19] extended this work in multiple ways to propose a transit evaluation framework. This framework consists of three stages. The first stage data pre-processing divides the urban area into equal size grids, which can be represented as a graph: with the grids as nodes and connected via edges that represent the road and transit system. The second part of the pre-processing consists of aggregating the bus-stops and trajectories into that grid system. The second stage consists of data-driven modelling, modelling the decision process as MDP and and derive decision making features from the network such as number of transfers, number of transit options,transit mode, travel time, fare, etc. The final stage is to use this work to learn about rewards, preferences and user choices in order to evaluate transit plans. In the study they describe a preference learning algorithm Inverse Reinforcement Learning with Suboptimal Policy (IRL+SP) that can capture non-linear reward functions of travelers. This algorithm works with the principle of maximum entropy and assumes that experts make decision with soft-max based sub-optimal policies.

To study how well IRL+SP performs in reward learning, Wu et al. [19] compare it to IRL and Apprenticeship Learning (AL) and claim that it leads to the lowest ridership vector difference, that IRL and IRL+SP converge faster than AL. To learn how well their algorithm performs in ridership prediction they combine IRL+SP with machine learning techniques such as random forest, lasso regression and linear regression. They compare this with a directly trained machine learning model and to a multinominal logit (MNL) model and claim to have the lowest prediction relative error. In this study they correctly note that MNL considers a route choice as a single decision of the entire trajectory instead of a sequence of decisions.

2.2 Discrete choice modeling of travel routes

Since the 1970's discrete choice modeling has been a leading method to understand choice behaviour of individuals in a wide range fields such as marketing, economics and transportation. Described by McFadden et al. [15] in 1973, discrete choice modeling has subsequently been extended over the decades in order to overcome specific limitations such as overlapping alternatives and correlations over time. The study of the specific field of route choice, is however more complicated than a choice between easily enumerable distinct alternatives, since route choice is typically not one single choice but instead a sequence of choices at each intersection, each transit stop, each mode, etc. This leads to very large choice set that is theoretically infinite due to loops. Often there can also be a large overlap between different route alternatives leading to difficulties for choice modeling. Two commonly used approaches are highlighted here.

A first approach to route choice modeling dates from the 1990's using a choice set consisting of the observed paths plus paths generated by a route choice generator. This approach has been used to estimate a number of different choice model forms such as multinominal logit and mixed logit. This approach comes with limitations: as discussed in Koch et al. [10], these route choice generators do not necessarily create realistic routes; and Frejinger et al. [7] argue that parameter estimates can vary significantly depending on the bias of the route choice generator. To address the specific issues with the overlap between difference alternative paths and the resulting correlations, multiple extensions have been proposed to attempt to avoid erroneous path probabilities and substitution patterns. The most two popular are path size logit Ben-Akiva and Bierlaire [1] and C-Logit Cascetta et al. [3], which decrease the utility of overlapping paths proportional to the overlap with other paths included in the choice set.

A second approach is to achieve a consistent choice set by sampling as proposed by Frejinger et al. [7]. This approach attempts to set up a sampling protocol in order to obtain unbiased parameter estimates from the route choice sets to neutralize the bias introduced by the route choice generator.

2.3 Dynamic discrete choice modeling of travel link sequences

An alternative approach uses link-based Markov decision process to model route choice as a series of sequential decisions. First proposed by Fosgerau et al. [6] it uses a linear system of equations to efficiently compute choice probabilities by using a solver to solve Bellman equations.

An incidence matrix is established that defines the exponential utility to perform action a from state k:

$$M_{ka} = \begin{Bmatrix} \delta\left(a|k\right) e^{\frac{1}{\mu} v(a|k)}, & a \in A\left(k\right) \\ 0 & \text{otherwise.} \end{Bmatrix} \tag{5}$$

The size of the incidence matrix is given by $|\widetilde{A}|$ describing the number of states A and the number of dummy links d representing termination states of destination. As the dummy links d have no successors, the row $k = d$ will be zero. Secondly Fosgerau et al. [6] define a vector z of size $|\widetilde{A}x1|$ vector where $z_k = e^{\frac{1}{\mu} V(K)}$ and a vector b of size $|\widetilde{A}x1|$ where $b_k = 0, k \neq d$ and $b_d = 1$. Now given the identity matrix I, Fosgerau et al. [6] write the linear equation:

$$z = Mz + b \iff (I - M)z = b \tag{6}$$

This system has a solution if $I - M$ is invertible, which might not be the case. As Fosgerau et al. [6] note this is highly dependent on the balance between the number of paths that connect the nodes in the network and the size of instantaneous utilities $\frac{1}{\mu} v(a|k)$. They note that this issue is particularly important to consider when estimating a model, as depending on the value of β, $I - M$ can be ill-conditioned or even singular. Fosgerau et al. [6] note that this limits the possible values of parameters, as when equation 6 does not yield a valid solution for at least one observation, the log likelihood function is not defined. They suggest to deal with this issue by starting at a feasible point (meaning a large enough magnitude in the parameters) and then being conservative in the initial step size of the line search algorithm at the price of an increased number of iterations.

Mai et al. [13] proposed a nested recursive logit that relaxes the independence from irrelevant alternatives property of the logit model by allowing scale parameters to be link specific. Zimmermann et al. [21] subsequently look at bicycle route choice problem in the city of Eugene, Oregon. By using 648 observations of bike trips collected from 103 users. They test a long list of 14 potential parameters: length; link constant to penalize paths with many constants; length interacted separately with upslope, medium traffic, heavy traffic, regional multi-use path, bicycle boulevard, bike lane; bridge; bridge interacted with bike facilities; no turn; no turn interacted with crossroad; left turn interacted with crossroad separately for medium traffic and for heavy traffic.

In Mai et al. [11] an improvement is proposed to Fosgerau et al. [6] by reducing the numbers of linear systems that need to be solved. By adding all observed destinations in vector b of size $|\widetilde{A}x|D||$ it becomes possible to solve the problem one iteration instead of solving the system for each destination separately, allowing for 30 times performance gain in their example. They use this performance gain to propose a mixed recursive logit, which allows for random taste variation by adding a random value to the utility function and running the model n draws each iteration to allow for a random variation. They perform a case study in two cities. First a car route choice model in the Swedish city of Borlänge, with 466 destinations, 1832 observations and a bicycle route choice model in Eugene, Oregon with 286 destinations with a unknown number of observations.

In de Freitas et al. [4], recursive logit is used to model inter-modal travel based on a static network that describes various connections in Zurich, Switzerland. The street network consists of 30,372 links and 13,828 nodes and the transit network consists of 10,298 transit links and 1585 nodes.

In de Moraes Ramos et al. [5], a network composed of 520 links and 200 nodes in is considered, using (nested) recursive logit to see how travel information affects route choice behaviour, and what is the impact of the travel time representation on the interpretation of parameter estimates and prediction accuracy.

3 CASE STUDY

3.1 Collecting data on bicycle movements

For this study we used the Dutch 2016 FietsTelweek ("Bicycle Counting Week") data set ([2]) that is available at their website and contains 282,796 unique trips. During 7 days between the 19th and 25th of September 2016 approximately 29,600 bicyclists volunteered to track their bicycle movements using a smartphone app. For this case study we limited the study to bicycle trips to and/or from the city of Amsterdam, Diemen, Amstelveen and Ouder-Amstel, leaving 29,684 trips by an unknown number of bicyclists as any personal identifying information was removed.

The application observing the participants ran in the background of the phone to collect the bicycle movements of all participants using the phone's GPS and acceleration sensors. The cyclists used their bike in a way as often seen in the Netherlands, using their bike as transportation from and to work, supermarket, school, friends, etc. For privacy reasons the resulting data was anonymized by the data provider before making it publicly available (i) by the removal of user information to make it impossible to trace multiple trips to a single person and (ii) by rounding of the trip departure time into one-hour bins to the nearest hour and (iii) removal of the random number between 0 and 400 meters from the start and the end of the trip to obfuscate the true origin and destination of each trip.

In prior research with the same specific sub-selection of the data, we found in Koch et al. [10] that bicyclists in Amsterdam often deviate from the shortest path, more than car drivers, indicating that there are different and possibly also more factors that have an effect on the routes bicyclists in Amsterdam take. In Koch et al. [10] we focused on the concept of route complexity: counting the number of locations where people deviate from the shortest path, in the interest of improving route choice generation techniques and potentially get more insight into the motivations for the route choice for bicyclists. In this study we explore other effects on route choice using different methodologies, without looking at route complexity or where people deviate from the shortest path. In future research we intend to combine both streams of work.

3.2 Environmental variables

To collect a set of variables that would reasonably impact route choice of bicyclists we collected and processed open data sources to compute various explanatory variables describing each route. First of all for each link in the network we include the length of that link as ***distance*** and if that link is a dedicated cycle-way, we include the length as ***oncycleway***. Additionally we have a variable ***traveltime*** based on the length and an estimated speed based on the GPS observations. To include data about the environment of each link we extracted information of data made openly available by the city of Amsterdam. Firstly we pulled potentially relevant variables from a geographical data-set with land-use zones. To combine the street-network with other relevant geographical data-sets, we cut each street link into small segments of 5 meters and determined the distance of that segments to a geographical feature in the land use data-set. The variable ***nearwater*** measures the distance of street situated close to water bodies such as the canals of Amsterdam, (small) lakes, rivers and other water bodies wider than 6 meters. To determine a preference for routes through parks and forests we did the same thing with the variable ***neargreen***, measuring distance of street situated within a 25 meter radius of 'green' land used for parks, forests and meadows.

To see if the vicinity of busy roads, a major source of noise and pollution, has any impact on route choice we used a data set with the noise contours map of road traffic in Amsterdam as shown in figure 1. This data-set is produced by a model that estimates the level of exposure to traffic noise in this map there are four noise levels with respectively at least 55, 60, 65 or 70 decibels of noise. The variables ***near55db***, ***near60db***, ***near65db*** and ***near70db*** represent the distance of the street passing through these exposure zones.

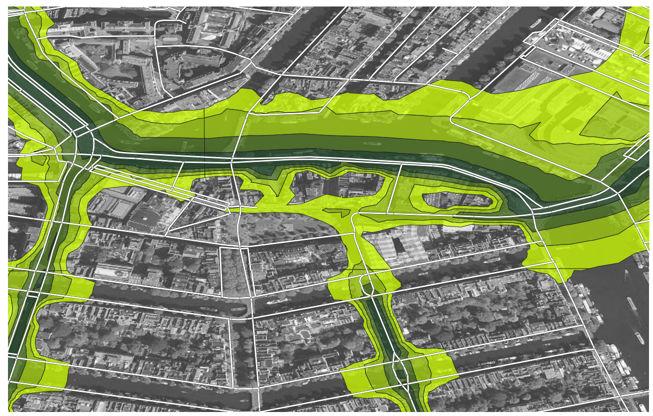

Figure 1: Noise contour map of Amsterdam, used for the variable that indicates the distance of a trajectory along roads with noisy traffic

4 RESULTS

Our initial attempt was to model the Amsterdam network with each intersection as a node and the streets as actions, following example in Zimmerman and Frejinger 2020. This resulted in a network with approximately 46,000 links and 30,000 observations, which we carefully controlled for full connectivity and no isolated graphs. Our motivation to model intersections as states instead of links as states was driven to lower the number of total states to be modeled, under the assumption that turn angles might have a low influence on bicycle route choice in Amsterdam. We tested the recursive logit model with the five variables *length*, *oncycleway*, *nearwater*, *neargreen* and *near55db*. However we were unable to get the solver to give plausible results for equation 5 as the solver would return incorrect results.

For the purpose of a better understanding of the algorithm we also implemented our own version of the original recursive logit model in Fosgerau et al. [6] and the significantly faster decomposition recursive logit model Mai et al. [11] in Python with SciPy and NumPy. Comparing the output of our Python implementation with publicly available MATLAB code and input data by authors of Mai et al. [13] a number of examples openly available, showed that the both implementations gave similar enough result values that the difference could be explained by floating point accuracy. This also meant that our Python re-implementation was similar ineffective giving plausible estimation results for the Amsterdam bicycle case.

Subsequently we simplified the study area to just the Amsterdam city center area containing only about 4500 links, excluding the entire municipality and surrounding suburbs. Again we carefully controlled for full connectivity and no isolated graphs. This too did not lead to plausible estimation results.

Based on the remark by Fosgerau et al. [6] on dense networks and the number of alternative paths, our next action was to simplify the street network in the Amsterdam city center and remove all footpaths to reduce the complexity of the network. Again we carefully controlled for full connectivity and no isolated graphs. We accordingly also removed all observations of GPS trajectories cycling over footpaths. This too did not lead to plausible estimation results.

Finally to transform our model to a model more similar to the studies in the literature, we instead created an edge-based network, instead of the intersection-based network. In the adapted implementation, each state is a street-segment and each action is a move to another street segment. This link-link approach allows the possibility to create new features with a boolean to indicate turns, left turns and u-turns, similar to the Borlänge model in Fosgerau et al. [6] and Mai et al. [11]. For the entire city of Amsterdam this model contains 40063 links as states with 137724 transitions between states; for the city center area only it consists of 4204 links as states with 15234 transitions.

With this network we were now able to solve the linear system to obtain a solution of z without (invalid) negative values for the entire city of Amsterdam. However even when setting the maximum number of links per observation at 30 links, we are still unable to calculate a log likelihood due to values of z_{origin} == 0 for one or more of the observations.

In the smaller area of the city center of Amsterdam it is possible to estimate a model but also with a relatively low limit of 30 links per observation, as a higher limit would again return zero values for z_{origin}. This meant we are able to process only 987 observations and 681 destinations. We listed the results of this model in table 1 where we would describe the betas to be plausible. An increase travel time would be a obvious cost. The negative value of an intersection, especially in the city centre where almost all road/bicycle intersections are equipped with traffic lights is also expected. The positive reward for for left-turn seems as expected to avoid crossing traffic. The positive reward for u-turn may seem odd, but u-turn costs on a bicycle should be less costly than in a motorized vehicle.

Subsequently we modelled travel-time with our 5 variables separately: 70db traffic noise distance, near green distance, near water distance, tree covered distance, cycle way distance. While all these models did converge, it was only able to invert the hessian to calculate a standard error for models: travel-time and length with tree cover; travel-time and length along water; travel-time and length with traffic noise. We listed these results in table 3. The $\beta length - treecover$ for is not significant, possibly because tree cover is less of an issue in the city center which is shielded by buildings. The result for traffic noise is not significant either. The only significant effect found besides travel-time was the distance travelled near water, possibly due to cycling along the cobble-stone paved narrow canals navigating between cars, trucks and tourists being perceived as disadvantageous to persons who cycle for daily activities.

Table 1: Results from estimated model on 987 observations in the city center of Amsterdam.

	value	std-error	t-test
$\beta_{travel-time}$	-2.9119	0.147249	-19.7751
$\beta_{intersection}$	-0.6356	0.045935	-13.8377
$\beta_{left-turn}$	-1.5717	0.075366	-20.8542
β_{u-turn}	0.4205	0.833094	7.5003
log likelihood	-4.489078		

Table 2: Results from estimated variable travel-time on 987 observations in the city center of Amsterdam.

	value	std-error	t-test
$\beta_{travel-time}$	-18.0368	0.006972	-2587.03
log likelihood	-10.279184		

Table 3: Results from 3 models (travel-time x length-noise, travel-time x length-water, travel-time x length-treecover) that were estimated on 987 observations in the city center of Amsterdam.

	value	std-error	t-test
$\beta_{travel-time}$	-18.03681	0.006972	-2587.035
$\beta_{length-noise}$	-1.9330	14.279491	-0.1354
log likelihood	-10.279184		
$\beta_{travel-time}$	-18.0369	0.00699	-2580.7
$\beta_{length-water}$	-2566.333	692.61447	-3.7053
log likelihood	-10.274874		
$\beta_{travel-time}$	-18.0368	0.006972	-2587.0
$\beta_{length-treecover}$	-1.9330	197.316315	-0.0098
log likelihood	-10.279184		

5 DISCUSSION

Given our experience with the Amsterdam model, we highlight several challenges during the estimation of the recursive logit model and reflect on why our initial plan for the model did not work out.

Negative reward formulation. In the original paper by Fosgerau et al. [6] on recursive logit it is mentioned that to formulate the path choice problem as a dynamic discrete choice model with the utility maximization problem consistent with a dynamic programming problem, the deterministic utility component is required to have negative value: $v_n(a|k) = v(x_{n,a|k;\beta}) < 0$.

As an experiment we set up a network based on a simple grid layout, with 625 intersections, allowing the user to move left, right, up, down. There is one diagonal connection across from the top left corner to the bottom right corner. We included each segment between the intersections as a single unit of distance. See figure 2

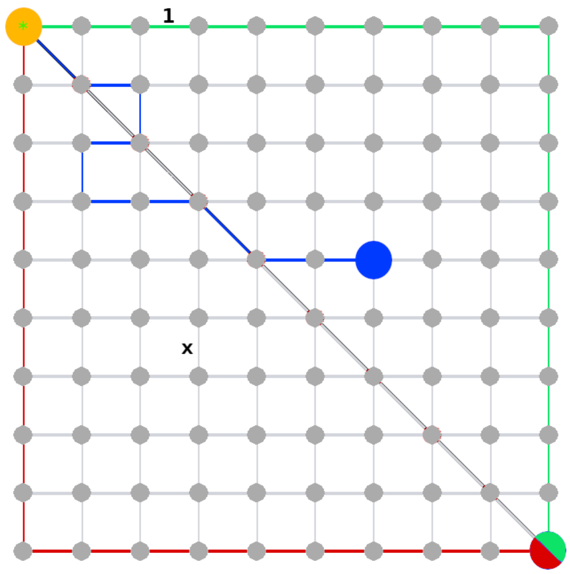

Figure 2: Two fixed paths to same destination along the boundaries of the graph (in red and green), plus example of one randomly generated path (in blue). All paths start at the top left corner and end respectively at the large red/green circle and the large blue circle.

for a visualization of 10 by 10 grid. We set up 4 different variables: $\beta_{distance}$ for the unit distance, $\beta_{intersection}$ that counts each intersection passed, β_{left} that counts each move towards the left side of the grid, $\beta_{diagonal}$ that counts each diagonal move. We included two observations across the top and right of the grid and a observation across the left and bottom of the grid and a series of 10 random observations that have a strong preference to move diagonally when possible. This model estimated with a log likelihood of -10 and $\beta_{distance} = -1.54467129, \beta_{intersection} = -2.04467129, \beta_{diagonal} = -2.09161539, \beta left = -81.34025$.

What we observed is that altering the attribute of a single link of this model to make the utility of that link positive lead to the inability of the linear solver to return a valid solution and thus not being able to find a log likelihood or estimate a model.

An implication that when using recursive logit you should aim for only including costs in your function u. In practise this might turn out tricky as cost variables may turn out to be correlated with reward variables not included in your model. For example heavy traffic near a bicycle path may seem like a cost variable at first, but as such roads are likely equipped with street lights in contrast to a path through a dark and empty park, such variable may turn out to have a negative cost.

Valid initial parameters and length of observations. To take a closer look at how difficult it can be to determine a valid initial parameter prior to iterative solution of the system, we proceeded to look at solely at travel-time without any other features in the model. To do so, we manually computed the log-likelihood function for a

Table 4: Descriptive statistics for variables length and travel-time in the city center of Amsterdam

	min	max	median	mean	std-dev	kurtosis
travel time	0.0102	21.3616	0.50330	0.8607	0.99338	25.9047
length	0.0007	0.51691	0.02760	0.0496	0.054914	5.70730

range of the $\beta_{travel-time}$ parameter in the range between -1 and -25. We saw that only in a small window of $\beta_{travel-time}$ between approximately -18.02 and -21.01 a valid log likelihood function exists. For a $\beta_{travel-time} <= -18$ the equation system would return an invalid sign for the log likelihood, for $\beta_{travel-time} >= 21.01$ at least one of the observations would return a $exp(V) = 0$ for at the starting value.

This narrow range was achieved with a number of links in each observations limited at 40. If we allowed observations with more links we were unable to find a window of initial parameters where the log likelihood function is valid at all. We see numerical issues as the root cause of this. As a long recursion will be a sum of each link utility, with high values due to the exponential, we expect these results to be caused by overflows and under flows in the solver.

The distribution of values of features and network degree centrality. Subsequently we attempted a similar experiment with the only feature in the model being β_{length}, which is correlated with $\beta_{travel-time}$. We were unable to find an exact parameter of β_{length} that is valid, but deduce it is somewhere between -413.6 and -413.7, based on where the solver returns a valid solution but $exp(V) = 0$. To look at the difference between both variables we will describe some statistics in table 4 and a histogram plot of length in figure 3 and travel time 4.

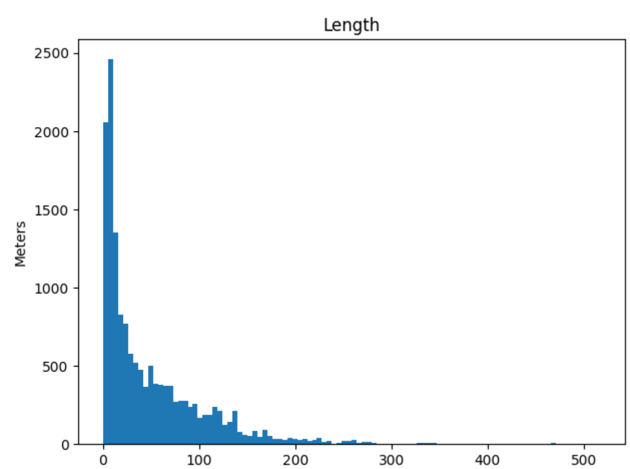

Figure 3: Histogram of the variable length of the bicycle network in the city center of Amsterdam

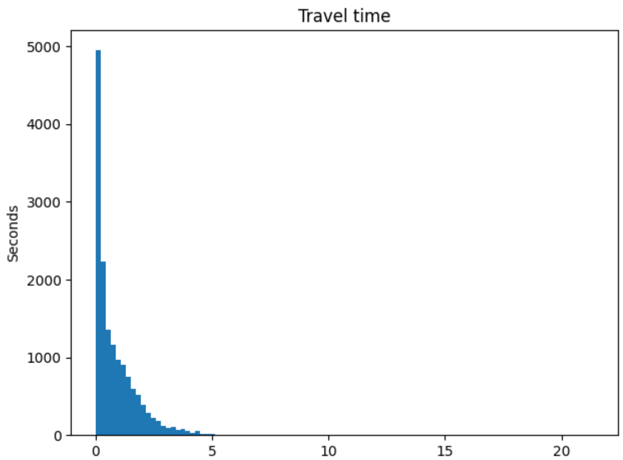

Figure 4: Histogram of the variable travel time of the bicycle network in the city center of Amsterdam

Based on the descriptive statistics we expect the same root cause that makes it difficult to find a valid starting parameter. The lower kurtosis in the distribution of length indicates a fatter right-tailed distribution presenting more possibility for a significant number of relatively large values to end up added together in the recursion on links. This too can lead to overflows and under flows making it difficult to find starting values betas due to numerical issues.

Another difference with existing studies in the literature that due to the complexity of bicycle infrastructure in Amsterdam, the number of possible options is higher than we would see in car route choice or in a city without two cycle-paths on both sides of major roads or two roads in both directions (for cyclists) along the canals.

6 CONCLUSION

Recursive logit is a promising solution for inverse reinforcement learning on specific route choice problems. However when designing your model and variables it is very important to keep the limitations of the linear equation system in mind. These limitations can make it impossible to estimate your model or lead to wrong estimations.

As recursive logit may fail to converge if even a single link has a (high) reward instead of cost, it is important to think through whether your variables are always costs for all links in the network. This can be hard in practice, as assumptions can be deceiving. For example you might model a bridge as a cost, as there is a small slope involved, however in reality people might prefer a route over a bridge as a form of sight seeing opportunity. Furthermore preferences can differ by person or vary over the time of day. For example a park might be a beneficial detour during the day, but during the night an empty badly lit park that feels unsafe might be worth a detour around instead. Better positio

7 FUTURE STUDY

For future study we are interested in the mechanics that lead to the invalid estimates by the solver when faced with numerical overflow and underflow issues. Could a more advanced solver resolve the issues we have seen? We are also looking into how well extensions of algorithms based on Maximum Entropy IRL of Ziebart et al. [20] will function with the Amsterdam bicycle network given the successful implementation of inverse reinforcement learning for bicycle paths in the work by Mo [16]. However given the limitations noted by Mo [16] on the same Amsterdam dataset, we expect challenges here as well.

REFERENCES

[1] Moshe Ben-Akiva and Michel Bierlaire. 1999. Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science*. Springer, 5–33.
[2] Bikeprint. 2017. Download bestanden Nationale Fietstelweek 2015 en 2016. http://www.bikeprint.nl/fietstelweek/
[3] Ennio Cascetta, Agostino Nuzzolo, Francesco Russo, and Antonino Vitetta. 1996. A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. In *Transportation and Traffic Theory. Proceedings of The 13th International Symposium On Transportation And Traffic Theory, Lyon, France, 24-26 July 1996.*
[4] Lucas Meyer de Freitas, Henrik Becker, Maëlle Zimmermann, and Kay W Axhausen. 2019. Modelling intermodal travel in Switzerland: A recursive logit approach. *Transportation Research Part A: Policy and Practice* 119 (2019), 200–213.
[5] Giselle de Moraes Ramos, Tien Mai, Winnie Daamen, Emma Frejinger, and SP Hoogendoorn. 2020. Route choice behaviour and travel information in a congested network: Static and dynamic recursive models. *Transportation Research Part C: Emerging Technologies* 114 (2020), 681–693.
[6] Mogens Fosgerau, Emma Frejinger, and Anders Karlstrom. 2013. A link based network route choice model with unrestricted choice set. *Transportation Research Part B* 56 (2013), 70–80. https://doi.org/10.1016/j.trb.2013.07.012
[7] Emma Frejinger, Michel Bierlaire, and Moshe Ben-Akiva. 2009. Sampling of alternatives for route choice modeling. *Transportation Research Part B: Methodological* 43, 10 (2009), 984–994.
[8] Tsubasa Hirakawa, Takayoshi Yamashita, Toru Tamaki, Hironobu Fujiyoshi, Yuta Umezu, Ichiro Takeuchi, Sakiko Matsumoto, and Ken Yoda. 2018. Can AI predict animal movements? Filling gaps in animal trajectories using inverse reinforcement learning. *Ecosphere* 9, 10 (2018).
[9] Edwin T Jaynes. 1957. Information theory and statistical mechanics. *Physical review* 106, 4 (1957), 620.
[10] Thomas Koch, Luk Knapen, and Elenna Dugundji. 2019. Path complexity and bicyclist route choice set quality assessment. *Personal and Ubiquitous Computing* (2019), 1–13.
[11] Tien Mai, Fabian Bastin, and Emma Frejinger. 2018. A decomposition method for estimating recursive logit based route choice models. *EURO Journal on Transportation and Logistics* 7, 3 (2018), 253–275.
[12] Tien Mai, Kennard Chan, and Patrick Jaillet. 2019. Generalized Maximum Causal Entropy for Inverse Reinforcement Learning. *arXiv preprint arXiv:1911.06928* (2019).
[13] Tien Mai, Mogens Fosgerau, and Emma Frejinger. 2015. A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological* 75 (2015), 100–112.
[14] Tien Mai, Quoc Phong Nguyen, Kian Hsiang Low, and Patrick Jaillet. 2019. Inverse Reinforcement Learning with Missing Data. *arXiv preprint arXiv:1911.06930* (2019).
[15] Daniel McFadden et al. 1973. Conditional logit analysis of qualitative choice behavior. (1973).
[16] Michael Mo. 2019. *Cyclist's route choice prediction with inverse reinforcement learning*. Master's thesis. Universiteit van Amsterdam, the Netherlands.
[17] Quoc Phong Nguyen, Bryan Kian Hsiang Low, and Patrick Jaillet. 2015. Inverse reinforcement learning with locally consistent reward functions. In *Advances in neural information processing systems*. 1747–1755.
[18] Guojun Wu, Yichen Ding, Yanhua Li, Jun Luo, Fan Zhang, and Jie Fu. 2017. Data-driven inverse learning of passenger preferences in urban public transits. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 5068–5073.
[19] Guojun Wu, Yanhua Li, Jie Bao, Yu Zheng, Jieping Ye, and Jun Luo. 2018. Human-centric urban transit evaluation and planning. In *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, 547–556.
[20] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. 2008. Maximum entropy inverse reinforcement learning.. In *Aaai*, Vol. 8. Chicago, IL, USA, 1433–1438.
[21] Maëlle Zimmermann, Tien Mai, and Emma Frejinger. 2017. Bike route choice modeling using GPS data without choice sets of paths. *Transportation research part C: emerging technologies* 75 (2017), 183–196.

Appendix

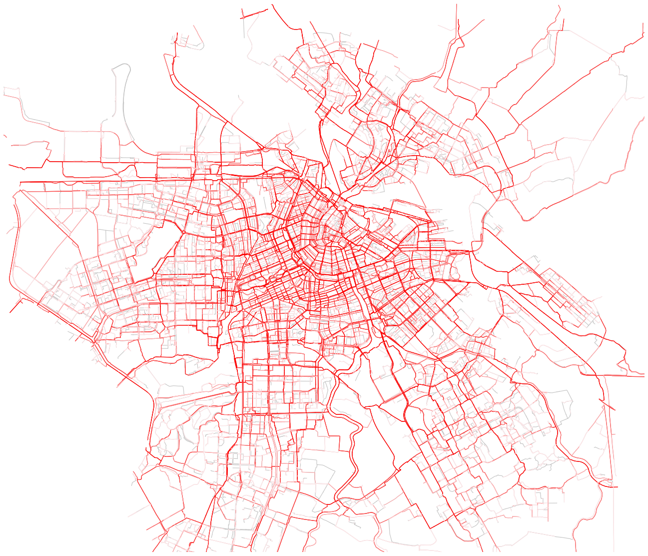

Figure 7: Map visualizing the trajectories observed in Amsterdam in the case study

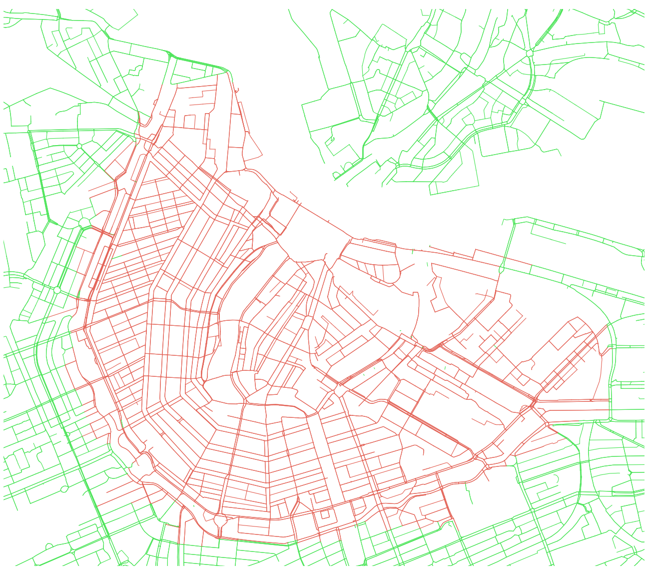

Figure 8: Map showing the selected area for city center of Amsterdam in red

Figure 5: Map visualizing the presence of the water near streets in Amsterdam

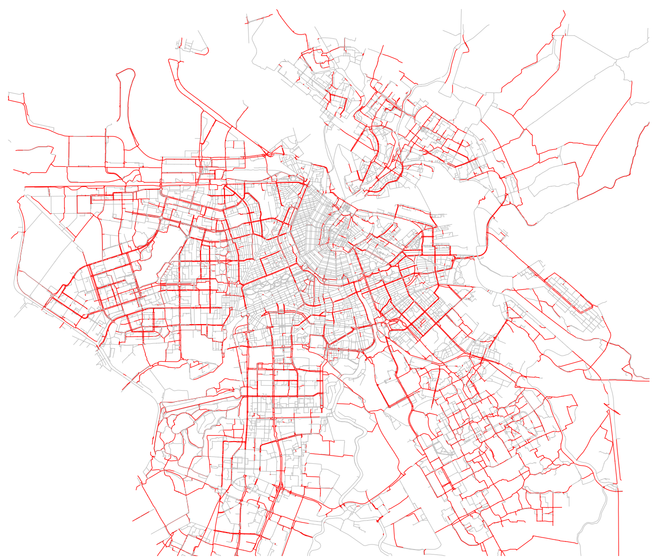

Figure 6: Map visualizing the presence of cycle-way infrastructure in Amsterdam