An Efficient Traffic Analysis and Optimization on the Dynamic Network Using Two Stage Routing Algorithm

A. Keerthika and K. Poonkavithai

Department of Computer Science and Engineering, MVIT-Pondicherry University, Pondicherry, India

Abstract

Now-a-days the transportation network traffic plays a vital role in the society. People are focusing on arriving at our destination as quickly as possible. With this lifestyle, people are not always aware of all the traffic conditions that are experienced while operating an automobile. In the existing system, reduction of green house gas emissions from transportation network based on road network graphs where all edges are annotated with accurate weight that capture environmental cost, fuel usage are needed for eco-routing. However, such weights are not readily available on the road network. So, randomly assign the weights for each road segments and it typically lack the speed pattern of the road network. This paper address these limitation, proposes two stage routing algorithm and weight propagation model to predict the cost of an arbitrary path on road network and accurately detects the traffic environment and also provides the optimal alternate route for destination.

Keywords: Transportation Networks, Weight Propagation Model, Vehicular Traffic, Routing

1. Introduction

Reduction in vehicular traffic is crucial in combating global environment change. To achieve these reductions, the transportation sector needs to achieve reductions. While improved vehicle and engine design are likely to yield vehicular emission reductions, eco-routing is readily deployable and is a simple yet effective approach to reducing emissions from road transportation.

Vehicle routing relies on a weighted-graph representation of the underlying road network. To achieve effective eco-routing, it is essential that accurate edge weights that capture environmental costs, e.g., fuel consumption or green house gas emissions, associated with traversing the edges are available. Given a graph with appropriate weights, eco-routes can be efficiently computed by existing routing algorithms. However, accurate weights that capture environmental impact are not always readily available for a road network.

This paper addresses the task of obtaining such weights for a road network from a collection of measured (trip, cost) pairs, where the cost can be any cost associated with a trip, e.g., green house gas emissions, fuel consumption, or travel time. Because the trips given in the input collection of pairs generally do not cover all edges of the road network and also do not cover all times of the day is a key problem. The cost of a trip, e.g., green house gas emissions, differs during peak versus off-peak hours. Thus, it is inappropriate to use costs associated with peak-hour trips for obtaining edge weights to be used for eco-routing during off-peak hours.
Considering the road network and trips shown in Figure 1, assume that the green house gas emissions of trip 1 (traversed from 7:30 to 7:33) and trip 2 (traversed from 23:15 to 23:17) are also given, and assume that we are interested in assigning green house gas emission weights to all edges in the network. The assignment of these weights to a large number of edges, e.g., BC, BD, EG, and FG, cannot be done directly since they are not covered by any trip.

However, for example, BD can be annotated by considering its neighbor road segment AB which is covered by trip 2. Assuming that the period from 6:00 to 8:00 is the sole peak-hour period (the remaining times being off-peak), trip 1 is not useful for assigning an off-peak weight to the edge AE because trip 1 traversed AE during peak hours. By taking into account the off-peak weights of IA and AB (covered by trip 2), it is, however, possible to obtain an off-peak weight for AE.

This paper proposes general techniques that take as input (i) a collection of (trip, cost) pairs, where trip captures the edges used and the times when the edges are traversed and the cost represents the cost of the entire trip; and (ii) an un-weighted graph model of the road network in which the trips occurred. In particular, this paper makes three contributions. First, a novel problem, road network weight annotation, is proposed and formalized. Second, a general framework for assigning time-varying trip cost based weights to the edges of the road network is presented, along with supportive models, including a directed, weighted graph model capable of capturing time-varying edge weights and a trip cost model based on time varying edge weights. Third, two novel and judiciously designed objective functions are proposed to contend with the data sparsity.

A weighted PageRank-based objective function aims to measure the variance of weights on road segments with similar traffic flows, and a second objective function aims to measure the weight difference on road segments that are directionally adjacent.

1.1. Overview

In this paper mainly focused on routing the vehicle which path is used several times in the transportation network, i.e., users preferred routes using weighted PageRank algorithm for eco-routing. The proposed system will enhance the driving direction based on two main features are one to detect accurate environment where the emission occurred and also the traffic strategies of the road segment using weight propagation model in the transportation network and another to find the optimal route to reach the destination at a time and low cost based on two stage routing algorithm. In these two stages are time dependent landmark graph model and
variance entropy based clustering model both are combined to provide the fastest route at low expenses.

2. Related Works

Over the past few decades, Motor vehicle traffic congestion is an everyday reality for millions of commuters. The ultimate goal in the transportation industry is to integrate detection, management, and control systems across large roadway networks and transmit pertinent information to travelers at any stage of a trip. The keystone of this integration is the ability to predict traffic conditions (i.e., flow, speed, and occupancy) on a road network. Trip cost estimation is a core component of our weight annotation solution. Given a set of (trip, cost) pairs as input, trip cost estimation aims to estimate the costs for trips that do not exist in the given input set.

Weight annotation can be regarded as a generalized version of trip cost estimation [1], since if pertinent weights can be assigned to a road network, the cost of any trip on the road network can be estimated. Most existing work on trip cost estimation [8-11] focuses on travel-time estimation. In other words, their work focuses on travel time as the trip cost. In general, the methods for estimating the travel times of trips can be classified into two categories: (i) segment models and (ii) trip models.

Segment models [10-13] concern travel time estimation for individual road segments. For example, observers (e.g., Bluetooth sensors or loop detectors deployed along road segments) monitor the traffic on road segments, recording the flows of vehicles along the road segments. Thus, travel-time estimation tends to concern particular road segments. For example, some studies model travel time on a particular road segment as a time series and apply autoregressive models [10] to estimate the travel time on the road segment. T-Drive [11] models time-dependent travel time distributions on road segments using sets of histograms and enables the inference of future travel times using Markov chains [14]. One study incorporates Lagrangian measurements [13] into existing traffic flow models for freeways to estimate travel time distributions on specific freeways.

In the intelligent transportation system research field [4, 15, 16], other travel costs (besides travel time) of trips are studied. For example, fuel consumption and GHG emissions of a trip can be computed based on instantaneous vehicle velocities and accelerations, the slopes of the road segments traversed, and the engine type. However, these methods are designed to estimate the costs of individual trips and are not readily applicable to the problem of annotating graph edges with trip cost based weights, notably edges that do not have any traversed trips.

2.1. Motivation

In the existing system is mainly focused on reducing the automobile emissions in the road network for effective eco-routing. Reduction in green house gas emissions from transportation calls for effective eco-routing, and road network graphs where all edges are annotated with accurate weights that capture environmental costs, e.g., fuel usage or green house gas emissions, are needed for eco-routing. However, such weights are not always readily available for a road network. This paper proposes a general framework that takes as input a collection of (trip, cost) pairs and assigns trip cost based weights to a graph representing a road network, where trip cost based weights may reflect green house gas emissions, fuel consumption, or travel time.

2.2. Topology of a Road Network

The topology of a road network is highly correlated with human movement flow including the movement of both pedestrians and vehicles. Edges with similar movement flows can be expected to have similar cost variables. Thus, if an edge is covered in TC, its
cost variable information can be propagated to the edges that have similar movement flow based similarity between edges using topological information of road networks.

A) Modeling Traffic Flows with PageRank

An idea of using PageRank for the modeling of web surfers to the modeling of vehicle movement in road networks. The original PageRank employs the hyperlink structure of the web to build a first-order Markov chain, where each web page corresponds to a state. The Markov chain is governed by a transition probability matrix $M$. If web page $i$ has a hyperlink pointing to web page $j$ then $M[i, j]$ is set to $1/\text{outDegree}(i)$; otherwise, it is set to 0. $M[i, j]$ indicates the probability of transition from state $i$ to state $j$. PageRank models a user browsing the web as a Markov process based on matrix $M$, and the final PageRank vector is the stationary distribution vector $x$ of matrix $M$. The PageRank of web page $i$, i.e., $x[i]$, indicates the probability that the user visits page $i$ or, equivalently, the fraction of time the user spends on page $i$ in the long run. The modeling movement of vehicles on a road network as stochastic processes is well studied in the transportation field. In particular, the modeling of vehicle movements as Markov processes is an easy-to-use and effective approach. Thus, the build a first-order Markov chain with a transition probability matrix derived from both the topology of the road network and the trips that occur in the road network. A state corresponds to an edge in the primal graph (i.e., a directed road segment), not a vertex (i.e., a road intersection). The PageRank value of a state indicates the probability that a vehicle travels on the edge or, equivalently, the fraction of time a vehicle spends on the edge in the long run. Thus, the PageRank value is expected to reflect the traffic flow on the edge.

Further, a series of topological metrics, including centrality-based metrics, small-world metrics, space-syntax metrics, and PageRank metrics, have been applied to capture human movement flows in urban environments. When using a graph representation of an urban environment, it is found that the classical and weighted PageRank metrics are highly correlated with human movements. Thus, if two edges have similar PageRank values, the traffic flow on the two segments should be similar.

When modeling web surfers, PageRank assumes that the Markov chain is time-homogeneous, meaning that the probability of transferring from page $i$ to page $j$ has the same fixed value at all times. In other words, matrix $M$ is static across time. In contrast, the time-homogenous assumption does not hold for vehicles traveling in road networks. For example, during peak hours, the transition probability from edge $i$ to edge $j$ may be substantially different from the probability during off-peak hours.

B) PageRank on Dual Graphs

PageRank was originally proposed to assign prestige to web pages in a web graph, where web pages are modeled as vertices and the hyper-links between web pages are modeled as edges.

![Figure 2. Dual Graph](attachment:image.png)
In Figure 2 assign PageRank values to edges, the primal graph $G = (V, E, L, F, H)$ is transformed into a dual graph $G^* = (V, E)$, where each vertex in $V$ corresponds to an edge in the primal graph, and where each edge in $E$, denoted by a pair of vertices in $V$, corresponds to a vertex in the primal graph. To avoid ambiguity, use the terms edge and vertex when referring to primal graphs and use dual edge and dual vertex when referring to dual graphs. Further, use the term weight when referring to the weight of an edge in a primal graph.

C) Weighted PageRank Computation

Definition of Dual Weights: In the original PageRank algorithm, a vertex propagates its PageRank value evenly to all its out-link neighbors. In other words, the dual weight for each dual edge from dual vertex $v_j$ is set uniformly to $1/|\text{OUT}(v_j)|$. The uniform weights on the web graph indicate that a web surfer chooses its next target web page without any preferences to continue its random surfing. However, in a road network, such non-preference surfing usually does not occur.

D) PageRank-Based Topological Constraint Objective Function

After obtaining the weighted PageRank values for every dual edge, the topological similarity between two edges in the primal graph is quantified. Based on the topological similarity, a PageRank-based Topological Constraint function is incorporated into the overall objective function. The intuition behind the constraint function is that for the same traffic category tag, if two edges have similar traffic flows, their cost variables tend to be similar as well.

E) Properties of PageRank on Road Networks

Web graphs and road network graphs are quite different, rendering it of interest to study the distributions of PageRank values on the two kinds of graphs are normalized (to (1, 100) PageRank values with respect to the percentage of vertices having the PageRank values, on a graph (WEB) representing a part of the Web and a dual graph (NJ) representing the road network.

Therefore, the PageRank values provides the important for top rank routes in the network but this weighted PageRank algorithm links cannot always be considered a recommendation and as noted earlier people are now trying to influence the links. So that ranking pages in network has been reported to be decreasing the importance.

The major weakness of PageRank’s focus on the importance of a page rather than on its relevancy given to the user query so rank of the page is closely related to the page visibility on the web, which in turn new one is not recognized first.

3. Proposed System

The proposed system to enhance the driving direction based on weight propagation model and two stage routing algorithm. Weight propagation model is used to determine the traffic status of the road networks where weights are calculated based on travel time, fuel emission and cost) to detects the traffic. The two stage routing algorithm is used for routing the vehicle based on time-dependent fastest route and traffic analysis based approaches for finding the optimal alternative route to reach the destination.
3.1. System Architecture

![System Architecture Diagram](image)

Figure 3. System Architecture

The overall system architecture Figure 3 shows the efficient and effective system for routing the vehicle. It describes the user initially to specify the source and destination location where he/she wants to go. Then, the source and destination location enters into the processing module it calculate the weights for all edges of the road segments using weight propagation model with the help of internet. Calculate overall average weight of the road network will detect the traffic status. By using two stage routing algorithm to provide the optimal alternate route from the traffic. Finally, display the best route from source and destination.

3.2. Two Stage Routing Algorithm

In these two stages are time dependent landmark graph model and variance entropy based clustering model both are combined to provide the fastest route at low expenses.

3.2.1. Time-Dependent Fastest Route Model

A landmark is one of the top-k road segments that are frequently traversed by taxi drivers according to the trajectory archive used the landmark to model the taxi drivers' intelligence is that: 1) The notion of landmarks follows the natural thinking pattern of people, and can give users a more understandable and memorable presentation of driving directions beyond detailed descriptions. 2) The sparseness and low-sampling-rate of the taxi trajectories do not support the speed estimation for each road segment while estimate the traveling time between two landmarks.

Meanwhile, the low-sampling-rate trajectories cannot over sufficient information for inferring the exact route traversed by a taxi. Here, to detect the top-k road segments as the landmarks instead of setting up a threshold, since a threshold will vary in the scale of taxi trajectories.
Time-Dependent Fastest Route Construction Algorithm

Time-dependent landmark graph to model the dynamic traffic pattern as well as the intelligence of experienced traffic control centre’s so as to provide a user with the practically fastest route to a given destination at a given departure time.

**Step1:** Road Segment. A road segment $r$ is a directed (one-way or bidirectional) edge that is associated with a direction symbol $(r:\text{dir})$, two terminal points $(r:s, r:e)$, and a list of intermediate points describing the segment using a polyline.

**Step2:** Dynamic Road Network. A dynamic road network $G_r$ is a directed graph, $G_r = (V_r; E_r)$, where $V_r$ is a set of nodes representing the terminal points of road segments, and $E_r$ is a set of edges denoting road segments.

**Step3:** Route. A route $R$ is a set of connected road segments, i.e., $R : r_1 ! r_2 ! \ldots ! r_n$, where $r_{k+1}:s = r_k:e$, $(1 \leq k < n)$. The start point and end point of a route can be represented as $R:s = r_1:s$ and $R:e = r_n:e$.

**Step4:** Taxi Trajectory. A taxi trajectory $T_r$ is a sequence of GPS points pertaining to one trip. Each point $p$ consists of a longitude, latitude and a time stamp $p:t$, i.e., $T_r : p_1 ! p_2 ! \ldots ! p_n$, where $0 < p_{i+1}:t_ - p_i:t < 4T$ $(1 \leq i < n)$. $4T$ defines the maximum sampling interval between two consecutive GPS points.

**Step5:** Landmark. A landmark is one of the top-k road segments that are frequently traversed by taxi drivers according to the trajectory archive.

**Step6:** Landmark Graph. A landmark graph $G_l = (V_l; E_l)$ is a directed graph that consists of a set of landmarks $V_l$ (conditioned by $k$) and a set of landmark edges $E_l$ conditioned by $t_{\text{max}}$.

3.2.2 Traffic Analysis based Approach

Traffic analysis based approach is devised to estimate the distribution of travel time between two landmarks in different time slots. A fast driving route saves not only the time of a traffic control centre but also energy consumption (as most gas is wasted in traffic jams). Therefore, this service is important for both end users and governments aiming to ease traffic problems and protect environment.

**Traffic Analysis Based Algorithm**

**Step 1:** First, sort according to the values of travel time $(t_1 - t_n)$, and then partition the sorted list $L$ into several sub-lists in a binary-recursive way.

**Step 2:** This step aims to split the x-axis into several time slots such that the travel times have a relatively stable distribution in each slot. It represent each travel time $y_i$ with the category it pertains to $(c(y_i))$, and then sort the pair collection $S_x = (x_i; c(y_i))|n i=1$ according to $x_i$ (arriving time).

**Step 3:** The optimism index indicates how fast a person would like to drive as compared to taxi drivers. The higher rank (position in taxi drivers), the faster the person would like to drive.

3.3. Weight Propagation Model

The goal of this task is to predict the cost for an arbitrary (possibly unknown) trajectory, based on a set of previous trajectory-cost pairs. The weight of an edge captures some cost associated with traversing the edge. It is a requirement to using a graph model for routing that all edges have weights. Weights that capture travel times and green house gas emissions can be extracted from GPS trajectory data collected from the network. GPS trajectory data typically needed to assign weights to all edges. In this equation
express the weights of all edges of the road segment. Where e is the link or edges of the road network, $C_e$ be the cost of the edges and $y(x)$ is the total weight network to detect the traffic jams.

A) Propagation Penalty

One of the most important features of networks is the fact that an event at one location can propagate to the neighboring locations through the connecting links. Thus, if a link $e$ has a significant deviation from its baseline state, then $e$'s neighboring links are expected to be influenced by that large deviation. This effect is easily understood by thinking about traffic jams.

B) Loss function

It is clear that using the baseline costs for such links is not a good solution, since such static information does not reflect any changes in the actual traffic flow. Predicting the cost of a trajectory including any unseen links appears to be difficult. However, introducing another assumption that calls weight propagation makes it possible.

4. Performance Evaluation

Considering that the travel time of a driving route depends on route, traffic and driver, it evaluates the following two aspects in the experiments. 1) How the proposed methods precisely predict the traffic conditions? 2) How these methods learn a user’s diver behavior accurately and estimate the travel time of a route for the user precisely? If the answers are effective then the system is valuable.

4.1. Evaluation on Traffic Prediction

Weight propagation mechanism is essential for trajectory cost prediction, 1) Prediction on a edge: taxi trajectories of the first two months as a training set and choose 12 days, consisting of 6 workdays and 6 weekends, from the trajectories of the third month as a test set. 6 out of the 12 days had normal traffic conditions, and the remainder had severe traffic conditions. The expectation of the travel times as a predictor calculated based on the inferred distribution. In [15] the ground truth of a given landmark edge is computed in terms of the average travel time of the transitions pertaining to the landmark edge in the time slot to be inferred. Many landmark edges with over several transitions are chosen for the evaluation. 2) Prediction on a road segment: test the performance of our method predicting traffic on road segments, using the tamilnadu traffic data.

Measurements: To quantify the accuracy of the traffic deduction using the weight propagation model has defined in [4]:

$$Y(x) = \sum_{e \in E} C_e$$

Where $e$ is the link or edges of the road network, $C_e$ be the cost of the edges and $y(x)$ is the total weight network to detect the traffic jams.

4.2. Evaluation on Routing

It is very difficult to directly evaluate whether a customized route for a real user is the actually fastest one due to the following two reasons. First, a user can only drive on one route at a given time. Second, it is not reasonable to request a single user to drive two different routes separately since the user can learn from their past driving experiences. So, the route driven later will benefit from the first test. Instead of directly finding the driving
route for a particular user by using two stages routing algorithm in [5] will provide fastest route.

To optimize driving routes [7, 8] has proposed optimal route providing services for a taxi driver by analyzing fleet trajectories. Here, they focus on taxi driving behavior in creating higher profit.

4.3. Efficiency

Besides being effective, the proposed system is also efficient due to the following reasons. First, weight propagation model calculate the weight of each road segments in the road network. So, the weight captures traffic status of all the subset of the original road network. This will speed up the traffic prediction on the road network. Second, the two stage routing algorithm used for finding the fastest route. Also, a rough route indicating the key directions reduces the search area on a road network and enables parallel computing when performing the detailed routing [11, 12]. So, the rough routing on the landmark is very fast.

5. Experimental Results

The result of this project it provides the optimal alternate path by using time dependent fast route is based on the time based road segments are chosen for routing this will be described in proposed time dependent fast route algorithm. The traffic analysis based routing approach will predict the road segment traffic in the transportation network.

Figure 4. Home Page- Route Specification

Figure 5. Possible Routes
6. Conclusion

In this paper describes the enhancement of driving direction to maintain traffic strategies on the road network. In the existing system will reduce the traffic
emission is made by assigning weight for all the edges of the network by using PageRank algorithm. It typically lacks the time consumption for preferring new routes, even if the route is of high quality is a serious problem in it. So that the proposed system, weight propagation model is used to detect the traffic environment and also provide the optimal alternative route from source to destination by using two stage routing algorithm. This will enhance the driving direction for effective eco-routing.

References