EcoSky: Reducing Vehicular Environmental Impact Through Eco-Routing

Chenjuan Guo, Bin Yang, Ove Andersen, Christian S. Jensen, Kristian Torp

Department of Computer Science, Aalborg University, Denmark
{cguo, byang, xcalibur, csj, torp}@cs.aau.dk

***Abstract*—Reduction in greenhouse gas emissions from transportation attracts increasing interest from governments, fleet managers, and individual drivers. Eco-routing, which enables drivers to use eco-friendly routes, is a simple and effective approach to reducing emissions from transportation. We present EcoSky, a system that annotates edges of a road network with time dependent and uncertain eco-weights using GPS data and that supports different types of eco-routing. *Basic eco-routing* returns the most eco-friendly routes; *skyline eco-routing* takes into account not only fuel consumption but also travel time and distance when computing eco-routes; and *personalized eco-routing* considers each driver's past behavior and accordingly suggests different routes to different drivers.**

I. Introduction

We are witnessing an increasing interest in green transportation from governments, fleet managers, and individual drivers alike [1]. To combat global climate change, governmental entities such as the EU, the G8, China, and Australia have set ambitious targets to reduce greenhouse gas (GHG) emissions from transportation. Individual drivers are increasingly aware of the possibility of using eco-friendly routes to reduce GHG emissions and, possibly, saving money at the same time.

Eco-routing is a simple and effective approach to reducing GHG emissions from vehicular transportation. Eco-routing relies on a weighted-graph representation of a road network, where each edge, or road segment, is associated with an eco-weight that captures the environmental impact (e.g., GHG emissions or fuel consumption) of traversing the edge. Studies suggest that eco-routing is able to achieve GHG reductions in the range 8–20% in varying settings [2].

The availability of accurate eco-weights is essential to enable eco-routing. Such weights are typically time-dependent and uncertain. For example, traversing an edge during peak hours may consume more fuel than during off-peak hours due to more frequent braking and accelerations. And under the same traffic conditions, aggressive driving may consume more fuel than moderate driving, resulting in significant variability and thus uncertainty. Further, the uncertainty itself may also vary over the time of day: it may be high during off-peak hours because drivers can drive as fast or slow as they want, while during peak hours, congestion may force drivers to drive similarly, thus reducing uncertainty.

These particular characteristics of eco-weights call for novel routing algorithms to enable practically useful eco-routing. First, eco-routing algorithms should support time-dependent and uncertain eco-weights. Second, practically useful eco-routing should not only consider the reduction of vehicular environmental impact, but should also consider other travel costs, such as travel time and distance. Optimizing for just a single of these is insufficient in practice. Third, it is desirable that eco-routing is personalized and takes into account individual drivers' particular driving behaviors and preferences.

Existing routing systems, e.g., PAROS [3] and EcoTour [4], primarily work on a road network that is annotated with time-homogeneous and deterministic weights and do not support personalized routing. Although PAROS takes into account multiple travel costs, it does not support time-dependent and uncertain weights.

We present a novel eco-routing system called EcoSky that consolidates techniques we have developed for eco-weight annotation and eco-routing. To the best of our knowledge, EcoSky is the first system that annotates a road network with time-dependent and uncertain eco-weights that are determined primarily based on GPS data and that offers different types of eco-routing based on eco-weights. We are not aware of other systems capable of offering such functionalities. An overview of EcoSky is presented in Fig. 1. Details of the modules are provided in Sections II and III.

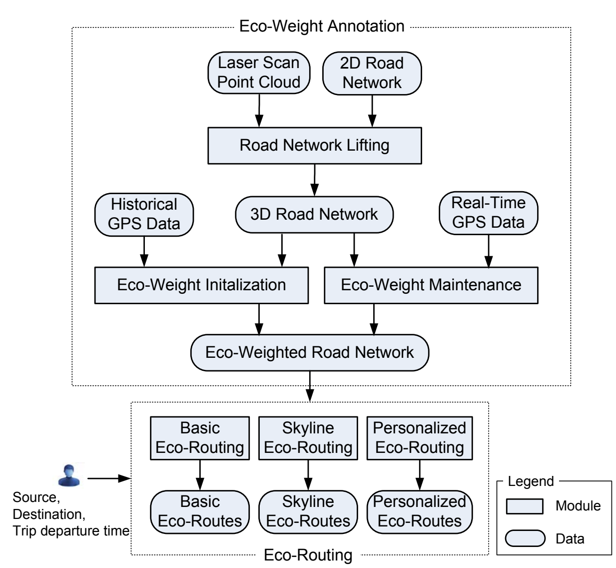


Fig. 1. EcoSky Overview

II. Eco-Weight Annotation

The eco-weight annotation module employs GPS data to annotate the edges of a road network with time-dependent and uncertain eco-weights, thus obtaining an eco-weighted road network. In this demonstration, eco-weights model fuel consumption.

An **eco-weighted road network** is a directed, weighted graph $G = (\mathbb{V}, \mathbb{E}, F)$, where $\mathbb{V}$ and $\mathbb{E} \subseteq \mathbb{V} \times \mathbb{V}$ is a vertex set and an edge set, respectively. A vertex $v_i \in \mathbb{V}$ models a road intersection or an end of a road. An edge $e_k \in \mathbb{E}$ models a directed road segment. Function F takes as input an edge e_k and returns its time-dependent and uncertain eco-weight.

The module contains three sub-modules—*road network lifting*, *eco-weight initialization*, and *eco-weight maintenance*.

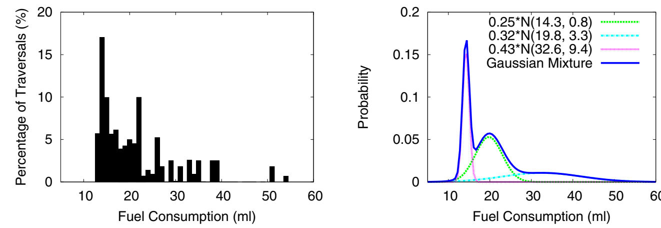
Road Network Lifting: This sub-module provides an efficient and accurate method for generating grades (i.e., incline/decline) of road segments. The grade of a road segment affects fuel consumption [2], but it is not available in digital maps such as OpenStreetMap. Our method is the first that lifts a 2D (i.e., latitude-longitude) road network to achieve a 3D (i.e., latitude-longitude-altitude) road network using an aerial laser scan point cloud, thus obtaining grade information on all road segments. It yields a 3D road network with a higher accuracy than what can be obtained using existing techniques. The algorithmic details are covered in our previous work [5].

Eco-Weight Initialization: This sub-module initializes the time-dependent and uncertain eco-weights based on historical GPS trajectories. Specifically, it annotates an edge with a set of (*interval*, *random variable*) pairs.

GPS trajectories are first split into *sub-trajectories*, where each sub-trajectory contains a set of GPS records that relate to a single edge. Next, since the GPS data does not contain fuel consumption information, we apply a vehicular environmental impact model [6] to estimate the fuel consumption based on vehicle velocities and accelerations (from the GPS data) and road grades (from the 3D road network). This yields a set of sub-trajectories, each of which, say *st*, has information about the fuel consumption of a traversal of an edge along with the traversal start time: $st = (edge, start\_time, fuel)$.

We divide a day into 96 15-minute intervals and distinguish weekdays from weekend days. When computing eco-weights for an edge e_i and for time interval I_j, we only consider the sub-trajectories that traversed the edge during the interval, i.e., $\{st | st.edge = e_i \wedge st.start\_time \in I_j\}$. Using these sub-trajectories, we derive a fuel-consumption histogram. An example is shown in Fig. 2(a), where the highest bar indicates that 17% of the traversals consume 14 ml of fuel during the particular 15-minute interval of interest.

Next, we represent the histogram by a Gaussian Mixture Model (GMM) based random variable. A GMM is a weighted sum of $K (K \geqslant 1)$ Gaussian distributions that is able to approximate arbitrary distributions. The GMM describing the fuel consumption distribution shown in Fig. 2(a) is presented in Fig. 2(b). The details of estimating a GMM given a collection of fuel consumption values can be found in our previous work [7].



(a) Traversals vs. Fuel Consumption (b) Learned GMM

Fig. 2. An Example for Learning GMM

Dealing with sparse data: To enable eco-routing, all edges must have eco-weights. However, only some edges are covered by GPS data during some periods. We deal with this sparsity using three strategies. We consider one edge at a time.

(i) When the number of sub-trajectories in a time interval is insufficient, the derived GMM may be over-fitting to the small number of available fuel consumption values. To avoid this, we consider sub-trajectories from nearby time intervals.

(ii) If there is a lack of time intervals with sufficient sub-trajectories, we derive a single GMM for all intervals using all sub-trajectories that occurred on the edge.

(iii) If an edge is not covered by GPS data, we assign a single deterministic value to the edge until sufficient trajectories can be obtained. This single value is inferred from those edges covered by GPS data that are adjacent to the edge or topologically similar to the edge [8]. If no such edges exist, we derive the value using the speed limit.

Eco-Weight Maintenance: This sub-module increases the accuracy of eco-weights by predicting the near-future eco-weights as real-time GPS data streams into the system. We propose a method for constructing a Spatio-Temporal Hidden Markov Model (STHMM) from sparse historical trajectories to model the spatio-temporal correlations of fuel consumption among all edges [9]. When real-time GPS records are collected on some edges, the STHMM is able to predict the near-future fuel consumption for the edges and their adjacent edges.

Summary: The novelty of the eco-weight annotation module lies in the following aspects: (i) lifting a 2D road network to a 3D road network using an aerial laser scan point cloud, (ii) learning arbitrary distributions from GPS data as time-dependent, uncertain eco-weights while dealing with sparsity, and (iii) maintaining the eco-weights as real-time GPS data streams in while taking into account data sparsity, correlation, and heterogeneity. We are not aware of existing techniques or systems that address any of these aspects or their combination.

III. Eco-Routing

We support three types of eco-routing—*basic eco-routing*, *skyline eco-routing*, and *personalized eco-routing*.

A route $\mathscr{R} = \langle e_1, e_2, \ldots, e_p \rangle$, where $p \geqslant 1$, is a sequence of edges, where $e_i \in G.\mathbb{E}$, $e_i \neq e_j$ if $i \neq j$, and where consecutive edges share a vertex. EcoSky takes as input a source, a destination, and a departure time, and it computes one or multiple routes depending on the type of routing service.

Basic Eco-Routing: Basic eco-routing aims to return the most eco-friendly route to users. The fuel consumption of an edge in an interval is derived as the expectation of the GMM random variable on the edge in the interval. We apply a time-dependent variant of Dijkstra's algorithm in this case.

Skyline Eco-Routing: Some eco-routing scenarios call for the consideration of multiple travel costs, not only fuel consumption. For example, FlexDanmark, a large fleet manager in Denmark, is considering the use of fuel consumption in addition to travel times and distances when scheduling trips.

In EcoSky, we consider three travel costs that are most relevant to eco-routing—fuel consumption (*FC*), travel time (*TT*), and distance (*DI*). We propose a natural notion of a stochastic skyline route for a given source-destination pair and a departure time. A stochastic skyline route is a pareto-optimal route with the property that no other route is better when considering all travel costs of interest.

We use random variable $FC(\mathscr{R}_i,t)$ ($TT(\mathscr{R}_i,t)$ or $DI(\mathscr{R}_i,t)$) to denote the distribution of fuel consumption (travel time or distance) of traversing route $\mathscr{R}_i$ starting at time t. Intuitively, the random variable $FC(\mathscr{R}_i,t)$ is the convolution of pertinent fuel consumption random variables of the edges in $\mathscr{R}_i$. Like eco-weights, time weights are time-dependent and uncertain and can be obtained using techniques similar to those used for obtaining eco-weights; and using time weights, the travel time random variable $TT(\mathscr{R}_i,t)$ can also be obtained using techniques similar to those used for obtaining $FC(\mathscr{R}_i,t)$. The distance random variable $DI(\mathscr{R}_i,t)$ is a deterministic value equal to the sum of the lengths of the edges on $\mathscr{R}_i$.

Given two routes $\mathscr{R}_i$ and $\mathscr{R}_j$ that connect the same source-destination pair, we say that $\mathscr{R}_i$ *dominates* $\mathscr{R}_j$ if for every $X \in \{DI,TT,FC\}$, $X(\mathscr{R}_i,t)$ is not stochastically dominated by $X(\mathscr{R}_j,t)$ and for at least one $X \in \{DI,TT,FC\}$, $X(\mathscr{R}_i,t)$ stochastically dominates $X(\mathscr{R}_j,t)$.

Based on the above, given a source v_s, a destination v_d, and a departure time t, skyline eco-routing $SER(v_s,v_d,t)$ computes the set of all routes such that no other route exists that dominates one of these routes.

$$SER(v_s,v_d,t) = \{\mathscr{R}_i | \neg\exists\mathscr{R}_j\ (\mathscr{R}_j \text{ dominates } \mathscr{R}_i)\}$$

Due to the space limitation, we refer the reader to our previous work for the technical details [7].

Personalized Eco-Routing: While skyline eco-routing is more comprehensive than basic eco-routing, it has the limitation that it does not take into account individual drivers' preferences and always suggests the same routes to all drivers.

We propose a method for effectively and automatically identifying context-aware driving preferences of a driver from historical GPS trajectory data of the driver while considering multiple, time-dependent, and uncertain weights. This allows personalized eco-routing to suggest routes according to each driver's preferences in different contexts [10]. For example, a driver may get the fastest route during peak hours, while getting a route that considers a trade-off between travel time and fuel consumption during a weekend trip. We capture a driver's preferences using a vector $\mathbf{w}$ that specifies the importance of different travel costs such as distance, travel time, and fuel consumption.

Next, a route $\mathscr{R}_i$ is assigned a cost vector $CV(\mathscr{R}_i) = (\overline{FC}(\mathscr{R}_i,t), \overline{TT}(\mathscr{R}_i,t), \overline{DI}(\mathscr{R}_i,t))$, where $\overline{X}(\mathscr{R}_i,t)$ is the expected value of random variable $X(\mathscr{R}_i,t)$.

Then, given a source v_s, a destination v_d, a departure time t, and a preference vector $\mathbf{w}$, personalized eco-routing $PER(v_s,v_d,t,\mathbf{w})$ computes the routes that have the least weighted sum of the expected costs w.r.t. $\mathbf{w}$:

$$PER(v_s,v_d,t,\mathbf{w}) = \{\mathscr{R}_i | \neg\exists\mathscr{R}_j(\mathbf{w}^{\mathrm{T}} \cdot CV(\mathscr{R}_j) < \mathbf{w}^{\mathrm{T}} \cdot CV(\mathscr{R}_i))\}$$

Summary: The novelty of the eco-routing module is two-fold: (i) it provides the first skyline routing service that takes into account multiple travel costs while considering fuel consumption and the combination of time-dependence and uncertainty; and (ii) it offers the first personalized routing service that automatically identifies context-aware driving preferences and uses the preferences to determine which routes to return.

IV. Demonstration Outline

EcoSky employs a GPS data set containing more than 180 million GPS records collected at 1 Hz (i.e., one per second) in Denmark in 2007 and 2008. EcoSky assigns time-dependent and uncertain eco-weights to more than 1,623K directed edges of the road network of Denmark that is obtained from OpenStreetMap (www.openstreetmap.org).

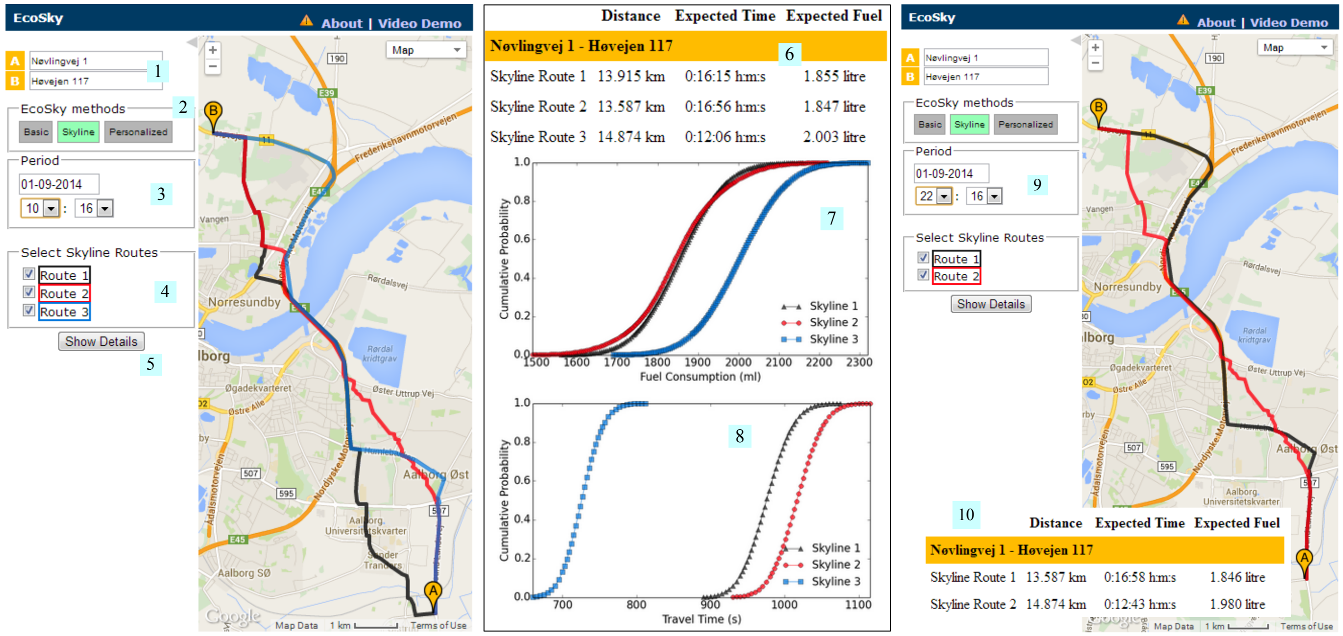
The lengths of edges are computed based on the coordinates of the corresponding vertices that are recorded in OpenStreetMap. The travel time of an edge is obtained as the difference between the times of the last and first GPS records of the sub-trajectories of the edge. The fuel consumption of an edge is computed based on the available GPS records of the sub-trajectories of the edge using the SIDRA-Running impact model [11]. Recent benchmarks [2], [6] indicate that this model is appropriate for this purpose.

EcoSky has a web interface (Fig. 3(a)) that includes a panel for entering and configuring queries and a map for visualizing eco-routes. In the panel, a source, a destination (Label 1, Fig. 3(a)), and a departure time (Label 3, Fig. 3(a)) can be specified by participants. The three buttons in the "EcoSky Methods" box allow participants to choose among basic, skyline, and personalized eco-routing (Label 2, Fig. 3(a)).

Skyline Eco-Routing Scenario: Participants can apply this functionality by clicking the "Skyline" button in the "EcoSky Methods" box. The skyline routes are then listed in the "Select Skyline Routes" (Label 4, Fig. 3(a)) and are shown on the map. Participants are able to check or uncheck the boxes in the "Select Skyline Routes" box to show the corresponding routes on the map or to hide them.

By clicking the "Show Details" button (Label 5, Fig. 3(a)), the statistics of the skyline eco-routes are presented in a pop-up window, as shown in Fig. 3(b). This functionality also applies to basic and personalized eco-routing.

The participants are shown the distance, expected travel time, and expected fuel consumption of each skyline eco-route (Label 6, Fig. 3(b)), as well as the cumulative distribution



(a) Skyline Eco-Routing Scenario-1 (b) Statistics (c) Skyline Eco-Routing Scenario-2

Fig. 3. EcoSky Web Interface

functions (CDFs) of the fuel consumption and travel time (Labels 7 and 8, Fig. 3(b)). Recall that skyline eco-routing takes into account not only fuel consumption, but also travel time and distance. Specifically, the skyline eco-routes shown in Fig. 3(a) contain the shortest route (Route 2), whose distance is 13.587 km, and the fastest route (Route 3), whose travel time random variable stochastically dominates those of all the other routes (Label 8, Fig. 3(b)). The fuel consumption random variables of Routes 1 and 2 do not stochastically dominate each other but dominate that of Route 3 (Label 7, Fig. 3(b)).

Different departure times can yield different sets of skyline eco-routes. To demonstrate this functionality, the participants can change the departure time, such as from 10:16 (Label 3, Fig. 3(a)) to 22:16 (Label 9, Fig. 3(c)) to obtain a different set of skyline eco-routes (Fig. 3(c)), as well as different expected travel time and fuel consumption distributions of the skyline eco-routes (Label 10, Fig. 3(c)).

Personalized Eco-Routing Scenario: Participants can choose to impersonate one of the drivers for whom we have learned driving preferences in advance. The personalized eco-routing returns the route that has the minimum weighted sum of travel costs w.r.t. the driver's preference vector.

Additional Demonstration Content: Participants will also learn details on how EcoSky derives and maintains time-dependent and uncertain eco-weights and how EcoSky conducts routing on a graph with time-dependent and uncertain edge weights [7], and they will receive information on how to derive individual drivers' driving preferences based on their historical trajectories [10].

V. Conclusion and Outlook

We demonstrate the EcoSky system that enables eco-routing, which reduces the environmental impact of vehicular transportation. It is of interest to explore further how to support continuous eco-routing where eco-routes change in real time due to eco-weight updates caused by incoming GPS data.

Acknowledgment

This work was supported by the EU FP7-ICT-2011-7 project REDUCTION (http://www.reduction-project.eu/), with project number 288254.

References

[1] C. Guo, C. S. Jensen, and B. Yang, "Towards total traffic awareness," *SIGMOD Record*, 43(3): 18–23, 2014.

[2] C. Guo, Y. Ma, B. Yang, C. S. Jensen, and M. Kaul, "Ecomark: evaluating models of vehicular environmental impact," in *SIGSPATIAL/GIS*, pp. 269–278, 2012.

[3] F. Graf, H. Kriegel, M. Renz and M. Schubert, "PAROS: pareto optimal route selection," in *SIGMOD Conference*, pp. 1199–1202, 2010.

[4] O. Andersen, C. S. Jensen, K. Torp and B. Yang. "EcoTour: reducing the environmental footprint of vehicles using eco-routes," in *MDM*, pp. 338–340, 2013.

[5] M. Kaul, B. Yang, and C. S. Jensen, "Building accurate 3D spatial networks to enable next generation intelligent transportation systems," in *MDM*, pp. 137–146, 2013.

[6] C. Guo, B. Yang, O. Andersen, C. S. Jensen and K. Torp, "EcoMark 2.0: empowering eco-routing with vehicular environmental models and actual vehicle fuel consumption data," *GeoInformatica*, 2014 (to appear).

[7] B. Yang, C. Guo, C. S. Jensen, M. Kaul, and S. Shang, "Stochastic skyline route planning under time-varying uncertainty," in *ICDE*, pp. 136-147, 2014.

[8] B. Yang, M. Kaul, and C. S. Jensen, "Using incomplete information for complete weight annotation of road networks," *TKDE*, 26(5): 1267-1279, 2014.

[9] B. Yang, C. Guo, and C. S. Jensen, "Travel cost inference from sparse, spatio-temporally correlated time series using Markov models," *PVLDB*, 6(9): 769–780, 2013.

[10] B. Yang, C. Guo, Y. Ma, and C. S. Jensen, "Towards personalized, context-aware routing," in submission.

[11] D. P. Bowyer, R. Akcelik, and D. Biggs, "Guide to fuel consumption analyses for urban traffic management," *Australian Road Research Board*, 1985.