Safety-aware Routing Using Machine Learning

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
Safety-aware Routing Using Machine Learning

ABSTRACT

Traditionally, vehicle routing is optimized for travel time or distance. However, there are contexts where routes that prioritize safety over other criteria are needed. This disclosure describes techniques that use publicly-available vehicle crash data and machine learning models trained on roadway characteristics, driver behavior, environmental factors, etc. to predict the likelihood of an accident on a given route. The historical data and the predictions made by the model are used as signals to rank recommended routes in a navigation system.

KEYWORDS

- Safety-aware routing
- Navigation
- Turn-by-turn navigation
- Vehicular routing
- Collision risk
- Accident data
- Collision data

BACKGROUND

Traditionally, vehicle routing is optimized for travel time or distance. However, there are contexts where routes that prioritize safety over other criteria are needed. Some approaches to safety-aware routing mine social media to discover and geo-locate crime activity. Users are routed around crime hotspots. Other approaches use crowd-sourced perceptions of public safety to similarly provide routes around areas perceived to be dangerous.
Existing approaches to safety-aware routing generally do not account for the behavior of other drivers on the road or prior knowledge of the roadway itself e.g., accident history, traffic volume, speed limits, etc.

DESCRIPTION

This disclosure describes techniques that use publicly-available crash data, e.g., as published by municipal or national safety authorities, and machine learning models trained on roadway characteristics, driver behavior, environmental factors, etc. to determine the risk of an accident on a given route and segments thereof. The historical data and the predictions made by the model are used as signals to rank recommended routes in a navigation system.

Fig. 1 illustrates the use of a trained machine learning model (102) to predict the probability of a collision (106) given features (104) such as:

- **Time**, e.g., time-of-day, day-of-week, month-of-year, day-of-year, day-of-month, year, etc.
- **Road features**, e.g., number of lanes; number of left-turning lanes; number of right-turning lanes; number of parking lanes; number of lanes of unknown type; road length;
road width; pedestrian facilities, e.g., crosswalk, sidewalk, median refuge, etc.; priority lanes; end-point type, e.g., stop sign, lane merger, traffic light, etc.; distance to end-point; bicycle lanes, bicycle safety features, or other bicycle facilities; road conditions; road type, e.g., surface road or elevated highway; road surface, e.g., paved, concrete, asphalted, gravel, dirt, etc.; road usage; barriers; elevation; etc. Road features can be obtained from data that includes maps, satellite or images that depict streets, or from a real-time scene analysis of the road. For example, with user permission, a real-time view of the road can be obtained from cameras or sensors aboard a vehicle.

- **Weather conditions**, e.g., temperature, dew point, barometric pressure, humidity level, wind-speed, rainfall, snowfall, etc.

- **Traffic conditions**, e.g., average speed, maximum permissible speed, volume, etc.

The ground truth, e.g., labeled data, used to train the machine learning model includes publicly available collision reports, e.g., as published by municipal or national safety authorities. During training, collision reports that geo-locate a collision by latitude and longitude can be snapped to the road segment.

The trained machine learning model can be, e.g., regression learning models, neural networks, etc. Example types of neural networks that can be used include gradient boosting tree model, long short-term memory (LSTM) neural networks, recurrent neural networks, convolutional neural networks, etc. Other machine learning models, e.g., support vector machines, random forests, etc., can also be used. Further, in addition to the above-listed features, collision reports themselves can be used as additional features to predict the risk associated with a particular route.
Fig. 2: Collision risk map

Fig. 2 illustrates an example collision risk map for a portion of a city (fictional city A) for a certain hour of the day, generated using the techniques of this disclosure. In a manner similar to the techniques used to optimize a route for travel time or distance, a map-routing algorithm computes a route between a user-specified origin and destination by optimizing the route for collision risk. The map-routing algorithm can also compute a route that is optimized jointly for collision risk, travel time, distance, or other criteria.

Collisions or other types of accidents are relatively rare events with a large variance in the time between accidents. Historical data pertaining to collisions in a given region is correspondingly sparse, with patterns of accidents relatively difficult to discern. The risk estimation techniques chosen to determine the collision risk map are well-calibrated to low-probability events.
Aside from safety-aware routing, the collision map generated can also be used to disseminate risk information to the public and to enable safety-related interventions by government authorities, e.g., the installation of safety measures such as barriers, warning lights, signs, etc. Additionally, with user permission to access user location data, a safety warning can be provided in a navigation system or map if a user is about to drive through a section of the road predicted to have a high risk of collision.

![Model generalization](image)

**Fig. 3: Model generalization**

The machine learning model described herein can generalize, e.g., it can determine a collision-risk map for a city or region for which accident data is not included in the training data for the model. This is illustrated in Fig. 3, in which collision-risk weighted routes are computed.
for an origin-destination pair in a different city (fictional city B) with a model that has been trained with collision reports from fictional city A (and/or other cities).

As illustrated, each computed route can optionally be annotated with travel time, distance, and collision risk. The annotations enable the user to choose a route that is appropriate for their current trip. In the illustrated case, the route with lower predicted risk uses highway segments rather than downtown and actually has lesser travel time, although it has a greater travel distance. While Fig. 3 shows a numerical value of collision risk, actual user interfaces may present users with route options with predicted safety ratings as well as recommendations.

The described techniques can be utilized in a navigation application that runs on a smartphone or other device, in an in-vehicle navigation system, etc. Users are provided with guidance that the prediction of route safety is based on historical data and may not be accurate in light of actual events on the roadways. Further, users are provided with choices to select from available routes, e.g., including routes for which the collision risk cannot be determined with adequate confidence.

Further to the descriptions above, a user is provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user’s environment, vehicle, driving and/or route preferences, or a user’s current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user’s identity may be treated so that no personally identifiable information can be determined for the user. The user’s location is determined and used for route computation with specific user
permission. Thus, the user has control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques that use publicly-available vehicle crash data and machine learning models trained on roadway characteristics, driver behavior, environmental factors, etc. to predict the likelihood of an accident on a given route. The historical data and the predictions made by the model are used as signals to rank recommended routes in a navigation system.

REFERENCES


