DEVELOPING AI-DRIVEN SAFE NAVIGATION TOOL

Subasish Das, Ph.D.

Texas A&M Transportation Institute | Texas State University

12/05/2023

OUTLINE

- Introduction
- Research Problem
- Research Gap
- Methodology
- Key Findings
- Conclusion

INTRODUCTION

- Navigation has a rich history believed to have originated with seafaring and later extended to land, aeronautics, and space navigation.
- In all cases, navigators determined their position relative to familiar locations or patterns, using methods such as dead reckoning or celestial navigation.
- As road signage limitations became apparent, navigation systems naturally gained popularity among drivers. The first car navigation system, capable of guiding drivers with the help of a rolling map, was introduced in 1910.

RESEARCH PROBLEM

- Route guidance systems (RGSs) have been considered driving assistance technologies since 1960.
- The goals of RGSs were limited to driver assistance and travel time reduction. Online or mobile navigation tools or RGSs such as Google Maps, Apple Maps, Waze, and MapQuest provide the shortest or fastest route between origin and destination locations, with some other options such as avoiding toll roads or avoiding freeways.
- None of these applications provide information about safety.

RESEARCH GAP

- The fastest route is often associated with disruptions such as entering or exiting ramps in a quick fashion.
- On the other hand, the shortest route can be associated with other nuisances such as narrow lanes, lack of lighting, and other poor geometric design features.
- There is a need for a safe navigation tool that can provide accurate measures of safety by applying prior historical data and other key associated features in artificial intelligence (AI)-driven algorithms.

LITERATURE REVIEW

Safety Consideration

- Route safety factors include crime risk, health risks, vehicle crashes, pedestrian/cyclist accidents, and HAZMAT transportation risk, guiding analyses for different user needs.
- Safety considerations vary for vehicles, pedestrians, cyclists, and public transport, focusing on factors like crash severity, walkability, lighting, and weather for personalized route safety.

LITERATURE REVIEW

Route-finding Algorithms

- Researchers have used deep reinforcement learning to find efficient routes based on factors like crime incidents, vehicle speed, and road conditions.
- The algorithms for safe route-finding can be classified based on their predictive/reactive nature, static/dynamic properties, and centralized/decentralized designs
  1. Predictive vs. Reactive Algorithms
  2. Static vs. Dynamic Algorithms
  3. Centralized vs. Decentralized Algorithms

LITERATURE REVIEW

Predictive vs. Reactive Algorithms

Reactive algorithms rely on observed data, while predictive algorithms use models to anticipate future conditions.

Static vs. Dynamic Algorithms

The route-finding algorithm can be divided into static and dynamic, depending on if the route-finding system reacts to real-time information.

Centralized vs. Decentralized Algorithms

Decentralized is where individual users make decisions to maximize their benefit, or centralized, where the aim is to optimize the benefit of all users or society.

LITERATURE REVIEW

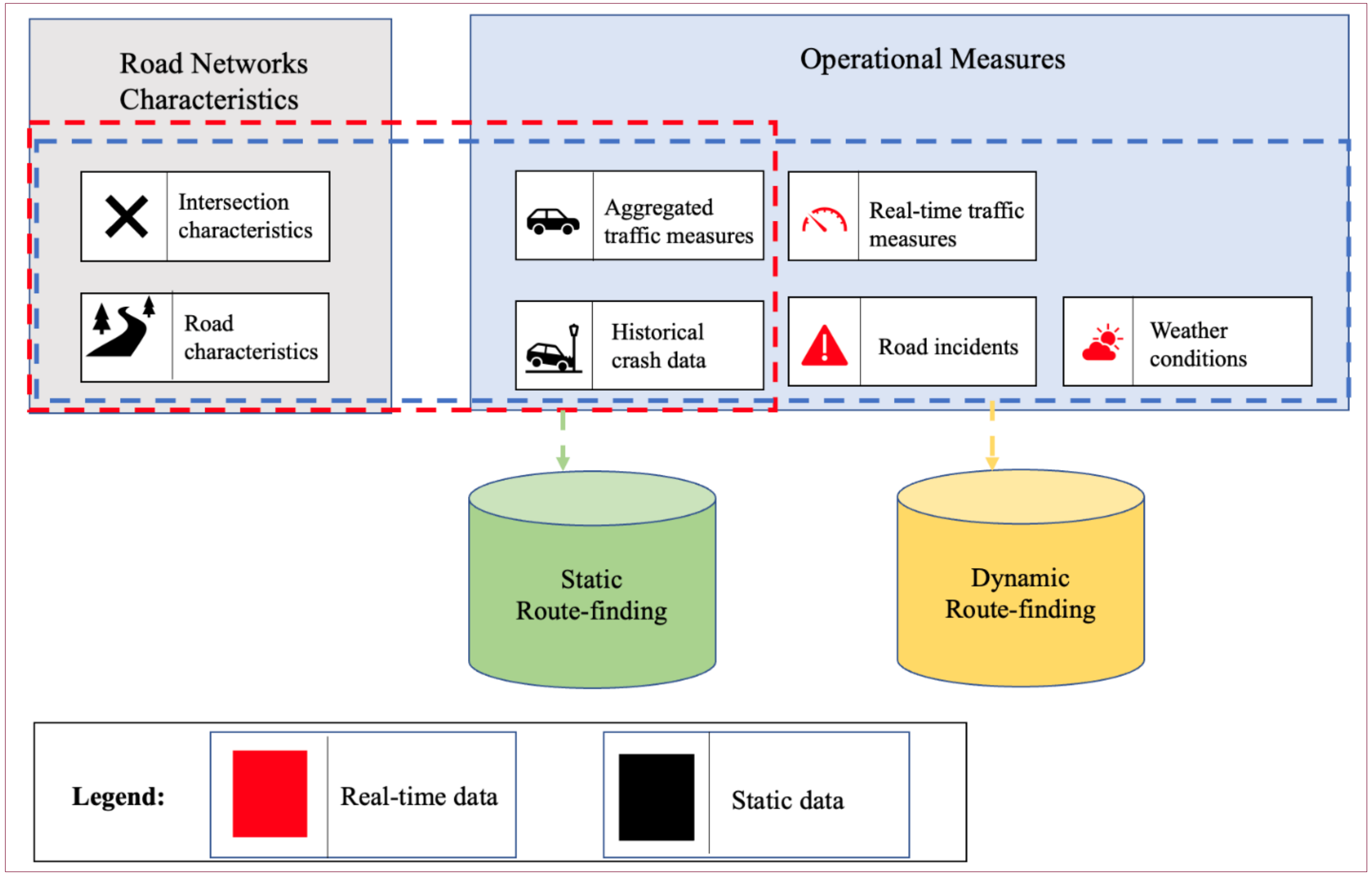
Safe Route-finding in Road Navigation

- Safe route-finding algorithms guide users by providing safety information on different routes.
- Relying solely on historical data to estimate crash risk has limitations because crashes are rare and involve factors that can change over time.
- Safe route-finding should consider these changing factors and offer future-oriented insights for initial route selection and alternative paths. Ideally, a dynamic, predictive algorithm is needed to provide reliable real-time crash risk predictions.

DATA NEEDS

- To develop a robust safe navigation tool, collecting comprehensive roadway inventory data alongside operational metrics like travel time, traffic volume, and incidents is vital.
- Utilizing diverse sources such as OpenStreetMap, private vendors like HERE, and multiple open-source datasets from DOTs and commercial providers such as Wejo ensures a well-rounded approach that integrates both public and private data for effective route planning.

DATA NEEDS


Road Networks Characteristics
Operational Measures
Intersection characteristics
Road characteristics
Aggregated traffic measures
Historical crash data
Real-time traffic measures
Road incidents
Weather conditions
Static Route-finding
Dynamic Route-finding
Legend:
Real-time data
Static data

DATABASES

Crash Record Information System (CRIS)

- The research team collected crash data from TxDOT for 5 years (2017–2021).
- There are 172 fields in the CRIS dataset.
- For example, in the 2021 CRIS data, among the 552,125 unique crashes, 542,445 crashes (~98.25%) contain absolute location data in the form of latitude and longitude coordinates.

DATABASES

Crash Record Information System (CRIS)

	2017	2018	2019	2020	2021
Total Entries	538,739	626,514	560,835	475,132	552,125
Entries with Lat-long	499,484	531,848	543,039	465,938	542,445
Entries Missing Lat-Long	39,225	94,666	17,796	9,194	9,680
Total Number of Variables	170	170	170	173	173

DATABASES

Roadway Highway Inventory Network Offload (RHiNO)

- The Roadway Inventory Annual Data published by the TxDOT is publicly available for download as a zip file that contains two shapefiles: one with assets and another with no assets.
- The primary key, all unique values, and the column of the shapefile without assets comprise the “GID” column.
- In 2021, there were 514,480 road segments recorded in the shapefile without assets.

DATABASES

Roadway Highway Inventory Network Offload (RHiNO)

Description	Count/Information
Total Records	514,480
Total Fields	7
Primary Key	GID
Entity	TxDOT
Source URL	https://www.txdot.gov/data-maps/roadway-inventory.html
Classification	Public Data
Datum	North American Datum 1983

DATABASES

North American Land Data Assimilation System (NLDAS)

- The goal is to construct quality-controlled and spatially and temporally consistent land-surface model (LSM) datasets from the best available observations and reanalyze the data to support modeling activities.
- Specifically, this system is intended to reduce errors in the stores of soil moisture and energy which are often present in numerical weather prediction models, and which degrade the accuracy of forecasts.

DATABASES

North American Land Data Assimilation System (NLDAS)

Description	Information
Spatial	0.125 ° x 0.125 °
Temporal	1 hour
Range	-108,25, -92,37
crs	+proj=longlat +R=6371200 +no_defs
Time	01/01/2017 - 12/31/2021
Source URL	https://disc.gsfc.nasa.gov/datasets?keywords=NLDAS&page=1

ROUTING ALGORITHMS

Google Directions API Shortest Path Algorithms

- Google Maps essentially uses two graph algorithms, Dijkstra's algorithm and an A* algorithm.
- Developers use API to display the shortest route in Google Maps.

Dijkstra's Algorithm

- It is an approach to solving the single-source shortest path problem for an assignment graph.
- This algorithm takes out the least distant of the unvisited nodes each time and updates the distances of the other nodes with that node.

ROUTING ALGORITHMS

A* Algorithm

- Considered as a best-first algorithm because each cell in the configuration space is evaluated by the Manhattan distance of the cell to the goal node and the length of the path from the initial node to the goal node through the selected sequence of cells.

ORS API Shortest Path Algorithms

- Algorithm is intended to minimize travel time rather than travel distance. In addition, the mode of transportation selected by the user and additional filters will affect which algorithm is used to calculate the route.

ROUTING ALGORITHMS

Contraction Hierarchies (CH) Algorithm

- ORS API uses the CH algorithm to calculate the shortest travel time route. Using this algorithm, roads are categorized ordinally by level with international highways being considered the highest level and residential roads being considered the lowest level.

Core-ALT (CALT) Algorithm

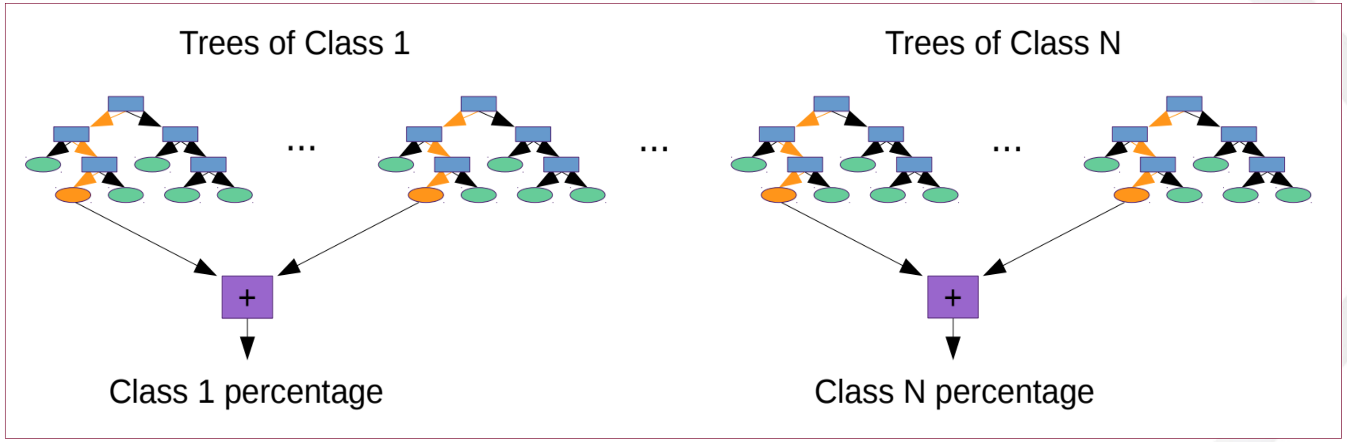
- When it is desirable to avoid obstructions such as rush-hour traffic jams, car accidents, or roads that are a higher safety risk to drive on, the CALT algorithm is used.

RISK SCORING

- Employed a comprehensive approach by utilizing four distinct AI models: random forest, gradient boosting, support vector regression, and CatBoost.
- Developed an effective risk-scoring system that considers various factors such as crash data, geometric properties of roadways, and weather information.
- CatBoost algorithm outperformed the other three algorithms in terms of predictive accuracy and overall performance. As a result, the final predictive values utilized in this study were derived from the CatBoost model.

CATBOOST

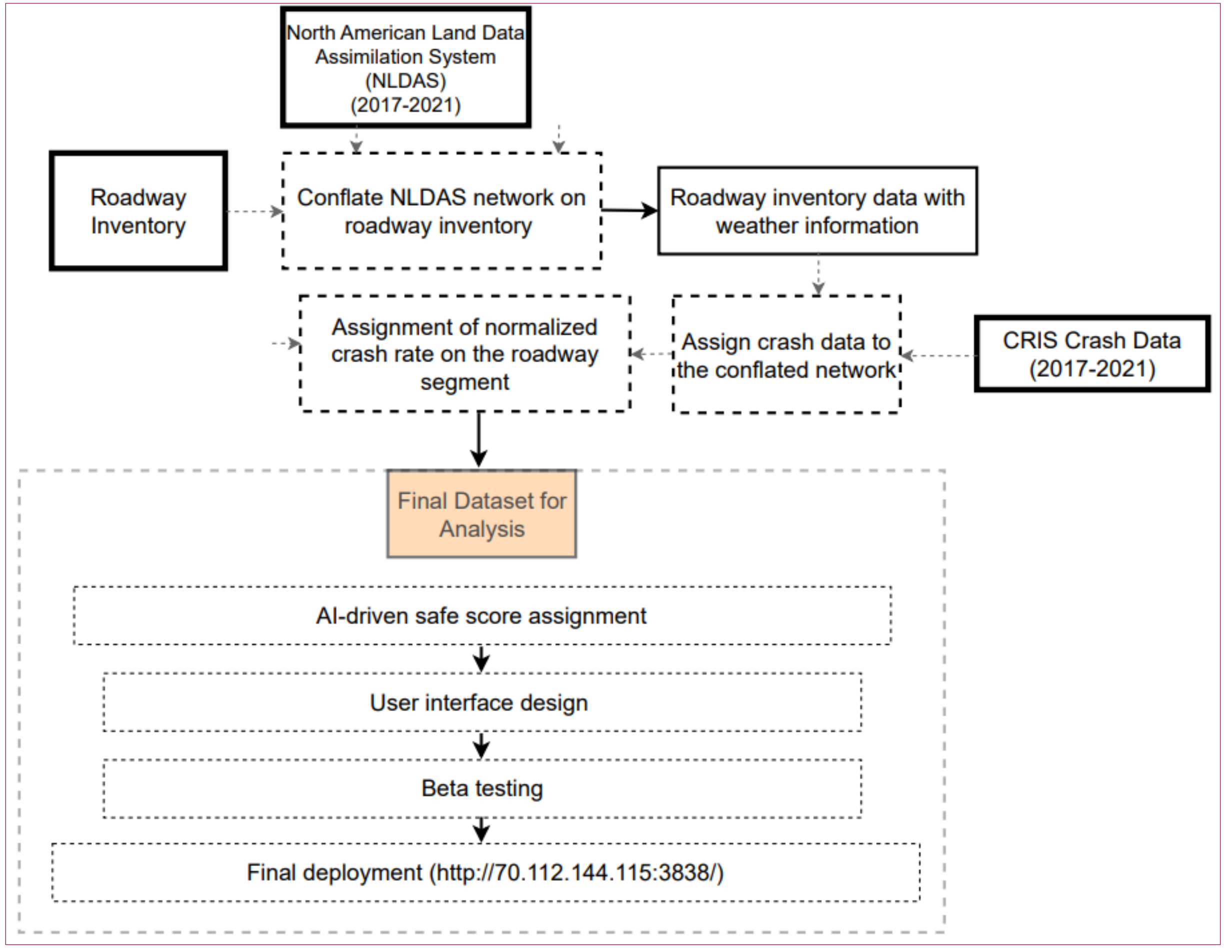
- Based on gradient boosting algorithm
- Handles missing data and numerical features
- Uses ordered boosting and random permutations
- Reduces overfitting and improves performance
- Automatically balances classes and selects learning rate



SAFE ROUTING TOOL

- The user interface of the developed system incorporates drop-down panels, allowing users to select their preferred temporal scope and prioritize annual durations for risk assessment.
- This user-friendly system design provides users with a comprehensive overview of the trade-offs between different routes based on safety considerations.
- By presenting both the visual representation of the safest route and a detailed tabulation of relevant metrics, users can make informed decisions that balance their priorities regarding travel time and road safety

SAFE ROUTING TOOL



SAFE ROUTING TOOL

Safe Navigation Tool

Consumer Mapping | Analyst Mapping | Statistical Analysis | UTC Project

Temporal Scope

Annual

Annual

All Years

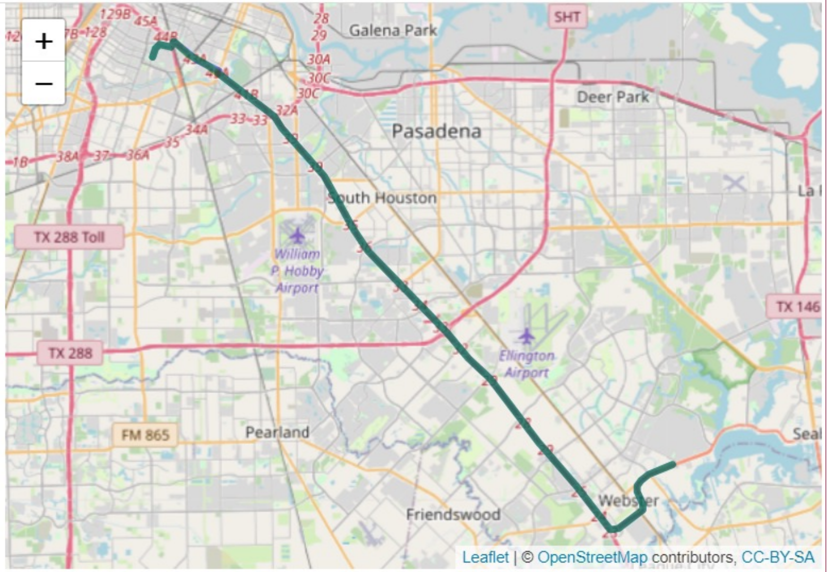
From

NASA, Houston

To

University of Houston

Directions



	Fast Route	Safe Route
Crash Score	8.65	6.88
Avg. Crash Score	0.07	0.06
High Crash Score	0.39	0.32
Road ID of High Score	8117	13388
Sum of Precipitation	0.07	0.07
Distance (miles)	23.27	23.39
Duration (minutes)	28.1	30.7

SAFE ROUTING TOOL

- The analysis of risk in road segments faces limitations due to the large variability in segment lengths.
- It is crucial to acknowledge that there might be localized clusters of high-risk areas. These clusters could exhibit more than 50 crashes within 100-foot sections along the 10,000-foot road segment.
- To address this limitation and reduce uncertainty, future research could consider dividing longer segments into shorter ones before associating crashes with their nearest segment.

KEY FINDINGS

- Accessing real-time data on traffic, road incidents, and weather conditions is crucial for dynamic and predictive algorithms in real-time routing.
- Limited public accessibility to real-time data poses challenges in developing effective route-finding algorithms, especially regarding traffic conditions and road closures.

KEY FINDINGS

- Crowdsourced data presents a potential alternative for obtaining real-time information, offering a solution to address the challenges of costly data collection for traffic and road conditions.
- Understanding the conflict between the fastest and safest routes influences user decision-making, impacting proactive initiatives for road safety education.

KEY FINDINGS

- Existing navigation apps lack real-time risk scoring, prompting the development of a tool using historical and real-time data sources, AI algorithms, and multiple factors for informed and safe route suggestions.

- Continued advancements are needed for a more robust safe navigation tool.

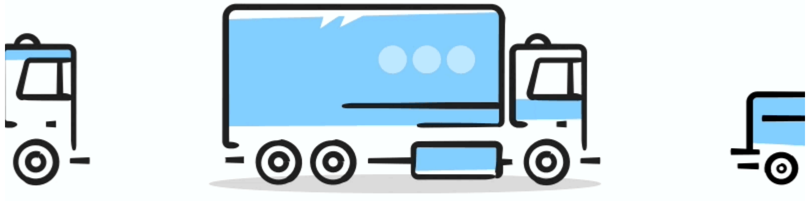
FUTURE RESEARCH

- Develop personalized vehicle-level crash prediction models to tailor trip details based on individual driving styles and historical crash records.
- Investigate short-term crash prediction models (e.g., daily, hourly) to integrate real-time safety insights into navigation tools.
- Introduce novel methods to assess overall route safety by aggregating risks at both road segment and vehicle levels.

FUTURE RESEARCH

- Explore leveraging crowdsourced data to support dynamic route-finding algorithms, especially for real-time crash and traffic information.
- Study user preferences regarding trade-offs between crash risks and travel time for informed decision-making.
- Compare centralized versus decentralized safe route-finding algorithms to understand their diverse impacts on road safety.

QUESTIONS?



Source: https://www.pinterest.com/pin/786723247204348 95/