School of Engineering and Applied Science (SEAS), Ahmedabad University

CSE 400: Fundamentals of Probability in Computing

Group-6 Intelligent Transportation System Scribe

Date of Submission: February 4 th, 2026

---

1 Project System and Objective

This project addresses a *probabilistic route selection problem* under certain uncertainty . Unlike deterministic shortest-path approaches that assume fixed travel times, this project models travel time as a random variable to capture the inherent unpredictability of real-world traffic conditions.

1.1 System Objective

The primary objective of the system is to **maximize the probability of arriving at a destination before calcualted time or a specified deadline** . Instead of minimizing average or expected travel time, the system focuses on identifying the *most reliable route* in terms of on-time arrival

Given a starting point s, a destination point d, and a maximum allowable travel time $T_{\max}$, the objective of the system is defined as

$$\max_{P \in \mathcal{P}(s,d)} \Pr(T_{\text{route}}(P) \leq T_{\max}),$$

where P denotes a feasible route from s to d, $\mathcal{P}(s,d)$ is the set of all such routes, and $T_{\text{route}}(P)$ is the random variable representing the total travel time along route P.

1.2 Project System Overview

The end-to-end system follows a probabilistic reasoning pipeline that integrates traffic data, uncertainty modeling, and reliability-based optimization:

- **Real-Time Traffic Data:** Live congestion indicators and sensor feeds are used to capture current traffic conditions and short-term variability.
- **Historical Congestion Data:** Past traffic patterns and delay statistics are utilized to estimate probability distributions of travel times.
- **Probabilistic Travel-Time Model:** Travel times on individual road segments are modeled as random variables, enabling explicit uncertainty modeling.
- **On-Time Arrival Probability Computation:** For each candidate route, the probability of reaching the destination within the specified deadline is computed based on the aggregated travel-time distribution.
- **Optimal Route Selection:** The route with the highest on-time arrival probability is selected as the optimal route.

1.3 Sources of Uncertainty

The probabilistic nature of the problem arises from multiple sources of uncertainty inherent in traffic systems:

- **Traffic Congestion Variability:** Fluctuating traffic demand and congestion levels cause random delays on road segments.
- **Accidents and Roadworks:** Non-recurrent events introduce unexpected and often significant delays.
- **Traffic Signals:** Signal timing and queue formation at intersections lead to stochastic waiting times.
- **Weather and Minor Incidents:** Weather conditions and small-scale disturbances further contribute to travel time variability.
- **Data Uncertainty and Model Approximation:** Travel-time estimation relies on historical and sensor data that may be noisy or incomplete, requiring simplifying modeling assumptions.

1.4 Problem Interpretation

By integrating real-time and historical traffic data into a probabilistic framework, the project formulates route selection as a reliability-driven decision-making problem. The system explicitly accounts for uncertainty and prioritizes on-time arrival, making it suitable for congestion-aware and deadline-sensitive routing scenarios.

2 Key Random Variables and Uncertainty Modeling

This section describes the probabilistic formulation of the routing problem by identifying the key random variables and explaining how uncertainty due to traffic congestion is modeled. The formulation captures variability at both the road-segment level and the route level, and directly supports reliability- based optimization.

2.1 Travel Time on Each Road Segment

Consider a route consisting of n road segments. The travel time required to traverse the i^{th} road segment is modeled as a random variable

$$T_i, \quad i = 1, 2, \ldots, n.$$

Each random variable T_i represents uncertainty caused by traffic congestion, signal delays, road conditions, and stochastic traffic flow. Instead of assigning a deterministic weight, each segment is associated with a probability distribution

$$P(T_i = t),$$

which may be estimated from historical traffic data or simulated traffic conditions.

2.2 Total Route Travel Time

The total travel time for a selected route is defined as the random variable

$$T_{\text{route}} = \sum_{i=1}^{n} T_i.$$

Since each segment travel time is uncertain, the total route travel time is also stochastic. Variability in T_{route} increases with route length and congestion intensity, making deterministic shortest-path approaches insufficient for reliable routing.

2.3 Arrival Time Relative to Deadline

To evaluate whether a route satisfies time constraints, an arrival indicator random variable is defined as

$$A = \begin{cases} 1, & \text{if } T_{\text{route}} \leq T_{\max}, \\ 0, & \text{otherwise,} \end{cases}$$

where $T_{\max}$ denotes the desired arrival deadline.

This variable indicates whether a selected route achieves on-time arrival. The probability $\Pr(A = 1)$ is used as the primary performance metric of the system.

2.4 Uncertainty Modeling and Probabilistic Assumptions

The following assumptions are made to model uncertainty in the routing problem: Initial Independence Assumption:

Conditional Dependence Between Segments:

Traffic Congestion Variability:

- **Initial Independence Assumption:** For the base model, travel times on individual road segments, represented as T_i are assumed to be independent random variables.
- **Conditional Dependence Between Segments:** Conditional dependence between segment travel times occurs when more than one road segment is impacted by a common traffic event, such as congestion spillback, traffic incidents, or signal coordination systems. In such cases, joint or conditional probability distributions are used to model the dependency.
- **Traffic Congestion Variability:** The probability distributions of segment travel times can change from one time interval to another to account for peak and off-peak traffic conditions.

2.5 Reliability-Based Objective

Based on the defined random variables, the routing problem is formulated as the maximization of on-time arrival probability:

$$\max \ \Pr(T_{\text{route}} \leq T_{\max}).$$

This formulation emphasizes reliability under uncertainty rather than minimization of expected travel time, making it suitable for congestion-aware and deadline-sensitive routing scenarios.

2.6 Modeling Assumptions

In order to model the uncertainty in travel times while keeping the model tractable, the following assumptions are made:

(a) **Stochastic Travel Times:** The travel time on each route segment is modeled as a random variable to represent the uncertainty introduced by traffic congestion, signal delays, and other unforeseen circumstances.

(b) **Baseline Independence Assumption:** To keep the model simple at this stage, the travel times on individual route segments T_i are modeled as independent random variables. This allows for easy aggregation of the uncertainty on individual segments to estimate the total route travel time.

(c) **Limited Dependency Modeling:** There could be some dependencies among the travel times on different route segments because of common traffic conditions. These dependencies are recognized but not modeled at this stage.

(d) **Stationary Traffic Conditions:** The traffic conditions are assumed to be statistically stationary for the duration of a route evaluation.

3 Probabilistic Reasoning and Dependencies section

This section is a overview how the probabilistic relationships are used in the project to support reasoning and decision-making under uncertain travel conditions.

3.1 Probabilistic Representation of Travel Time

- Travel time on each route segment is modeled as a random variable. This representation reflects uncertainty arising from varying traffic conditions.
- The probabilistic model enables the system to reason about a range of possible travel outcomes rather than assuming a single fixed travel time.

3.2 Route-Level Inference

- The total travel time of a route is obtained by aggregating the random travel times of its individual segments.
- This aggregation is used to compute the probability of arriving within a given time limit, which supports comparison and selection of candidate routes during decision-making.

3.3 Independence and Dependence Assumptions

- In the initial model, travel times of different route segments are assumed to be independent in order to simplify probabilistic reasoning and implementation.
- Possible dependence between adjacent segments due to shared traffic conditions is acknowledged but not explicitly modeled at the current stage.

Overall, probabilistic relationships are used to estimate on-time arrival probabilities and to identify routes that are more reliable under uncertainty.

4 Model–Implementation Alignment

This section explains how the probabilistic model is reflected in the current implementation and experimental setup, and how modeling assumptions influence design and evaluation choices.

4.1 Alignment Between Model and Implementation

- The probabilistic model defines travel time as a random variable, and the implementation follows this by treating route performance in terms of uncertainty rather than fixed values.
- The system evaluates candidate routes using the probability of arriving within a given time limit, which directly aligns with the probabilistic objective defined in the model.
- The implementation focuses on analysis and comparison of routes under uncertainty, rather than computing a single deterministic optimal route.

4.2 Modeling Assumptions Affecting Design Choices

- Independence between segment travel times is assumed in the current model to simplify probability estimation and reduce computational complexity.
- Traffic conditions are assumed to be stationary during a single route evaluation, allowing probabilities to be computed using fixed travel-time distributions.
- These assumptions make the early-stage implementation feasible and consistent with the probabilistic formulation, while acknowledging that they may not fully capture real-world behavior.
- The current design prioritizes clarity and interpretability of results over model completeness, which supports learning and incremental refinement in later milestones.

5 Cross-Milestone Consistency and Change

This section summarizes the current understanding of the project's probabilistic model and how it is expected to evolve across milestones.

5.1 Currently Well-Defined Components

- Travel time is treated as a random variable to represent uncertainty in the system.
- The system objective is defined using the probability of on-time arrival, which is used to compare candidate routes or solutions.
- At this stage, the project focuses on learning how probabilistic objectives can be analyzed rather than fully solved.
- The Genetic Algorithm is considered as an example to understand how guided optimization improves over a pure random approach.

5.2 Expected Refinements in Future Milestones

- Current assumptions, such as independence between uncertainty sources, are acknowledged as simplifications.
- Future milestones will focus on organizing assumptions more clearly and improving the analysis of probabilistic behavior.
- The Genetic Algorithm will be explored further to study its behavior under probabilistic objectives, with emphasis on analysis and reasoning.

Some modeling assumptions are recognized as simplifications that may require refinement. In particular, the assumption of independence between road segments is expected to be revisited as the group gains a deeper understanding of traffic dependencies.
Additionally, in future milestones we will try to incorporate improved handling of time-varying traffic conditions and more detailed uncertainty models as data availability and implementation maturity increase.

6 Open Issues and Responsibility Attribution

This section identifies unresolved probabilistic issues at the current stage of the project and outlines responsibility for addressing them in future milestones.

6.1 Open Probabilistic Issues

- Traffic congestion is uncertain and may change due to unexpected factors such as incidents or sudden demand variations, which are not fully captured in the current model.
- Travel time uncertainty is represented in a simplified manner, and the most appropriate probability distributions for different routes are not yet clearly established.
- The current formulation does not explicitly model dependence between adjacent route segments, even though shared traffic conditions may influence multiple segments.
- The impact of changing traffic conditions during route execution (non-stationarity) is acknowledged but not yet addressed.
- Uncertainty due to noisy or incomplete traffic data remains an open concern in estimating reliable on-time arrival probabilities.

6.2 Responsibility in Subsequent Milestones

- Data-focused tasks will work on improving data quality and gaining a better understanding of traffic variability and uncertainty sources.
- Probabilistic modeling tasks will focus on refining assumptions and improving the representation of travel-time uncertainty.
- Algorithm and system design tasks will analyze how uncertainty affects route selection and how probabilistic objectives can be handled more effectively.
- The evaluation tasks will focus on analyzing results and drawing insights rather than achieving a fully optimized solution.

Reference and paper we used for this project and especially for this millestone is:-

Andonov, G., & Yang, B. (2018). *A new formulation of the shortest path problem with on-time arrival reliability.* arXiv. https://arxiv.org/abs/1804.07829

End of Submission