PROTRIP: Probabilistic Risk-Aware Optimal Transit Planner

Pranay Thangeda and Melkior Ornik

Abstract— **Optimal routing in urban transit networks, where variable congestion levels often lead to stochastic travel times, is usually studied with the least expected travel time (LET) as the performance criteria under the assumption of travel time independence on different road segments. However, a LET path might be subjected to high variability of travel time and therefore might not be desirable to transit users seeking a predictable arrival time. Further, there exists a spatial correlation in urban travel times due to the cascading effect of congestion across the road network. In this work, we propose a methodology and a tool that, given an origin-destination pair, a travel time budget, and a measure of the passenger's tolerance for uncertainty, provide the optimal online route choice in a transit network by balancing the objectives of maximizing on-time arrival probability and minimizing expected travel time. Our framework takes into account the correlation between travel time of different edges along a route and updates downstream distributions by taking advantage of upstream real-time information. We demonstrate the utility and performance of our algorithm with the help of realistic numerical experiments conducted on a fixed-route bus system that serves the residents of the Champaign-Urbana metropolitan area.**

I. INTRODUCTION

Optimal passenger routing in transit networks is of significant practical importance and plays an essential role in improving passenger experience. Travel time in urban road networks is often stochastic as it depends on a number of factors, including but not limited to, traffic incidents, construction zones, and pedestrian movement patterns [1], [2]. Most of the smartphone-based navigation services generate the optimal path prior to departure by solving the trivial problem of minimizing the expected travel time using available data [3], [4]. Using least expected travel time (LET) as the metric for route selection in networks with stochastic travel times might not be always ideal as they have no consideration for the passengers' *delay tolerance*: their willingness to accept variability in the arrival time at their destination. Individuals have their own preferred delay tolerances, and even the same individual may have a different delay tolerance for different trip purposes. Therefore, inclusion of a metric that integrates the *route reliability*, i.e., the probability of arriving on time, into routing decisions is essential.

Given an origin-destination pair and a travel time budget, the reliability of the route can be defined as the probability of taking that route from origin to destination and realizing a travel time no greater than the given time budget [5]. Reliability-based optimal routing with stochastic travel times has been extensively studied and several solution algorithms exist in the literature for both the a priori variant that plans the entire route before departure [6], [7], and the adaptive variant that takes routing decisions online [8], [9]. While the reliability objective helps a passenger in planning the trip for on-time arrival tasks, it suffers the same drawback of the LET objective in that it is not suitable as a versatile routing solution. This paper proposes a multiobjective optimization model where the choice is based on both the reliability and the expected travel time of the route, with the passenger's delay tolerance specifying their relative priority.

In addition to only considering LET routing, much of the previous work on routing in road networks with stochastic travel times assumes independence of travel time between different road segments [7], [10]. However, this assumption often does not hold true in real-world traffic networks, as it becomes evident from several studies in the literature [11], [12]. While the problem of optimal routing with the LET objective in a network with correlated edge travel times has already been studied [13], [14], no significant work exists, to the authors' knowledge, on multiobjective optimal routing that considers the spatial correlation of travel times.

To address the above-discussed shortcomings of the existing transit routing systems, we propose PROTRIP: Probabilistic Risk-Aware Optimal Transit Planner, a tool and an underlying methodology to generate optimal routes in spatially correlated stochastic networks while considering multiple, potentially conflicting objectives. We handle the multiobjective model by assigning appropriate weights to scaled values of individual objectives. The weights, which decide the preference of one objective over another, are naturally expressed as a function of the delay tolerance provided by the passenger based on the passenger's preferences for a particular trip. A probabilistic model that explicitly models the random variables corresponding to the travel time on each road segment and the correlation between them supports all the decisions made by the tool. Given an origin-destination pair, a travel time budget, and a delay tolerance, PROTRIP provides an online routing strategy that provides the optimal route choice using the information obtained from the static schedule of the transit network, the real-time departure timings at different stops, and the probabilistic model incorporating spatial travel time correlations. We demonstrate the utility of our tool by implementing our algorithm to provide routing services in the transit system of the Champaign-Urbana metropolitan area in Illinois.

This work was supported by NASA's Space Technology Research Grants program for Early Stage Innovations under the grant 'Safety-Constrained and Efficient Learning for Resilient Autonomous Space Systems.'

Pranay Thangeda and Melkior Ornik are with the Department of Aerospace Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, USA. {pranayt2, mornik}@illinois.edu.

II. Preliminaries and Notation

We model a bus-based, fixed-route transit network using a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with n_v vertices and n_e edges. Each stop, defined as the location where a passenger can enter or exit the bus, is represented by a vertex $v_i \in \mathcal{V}, i \in \{1, 2, \dots, n_v\}$, and the road segment between any two stops in the transit network is represented by an edge $e_j \in \mathcal{E}, j \in \{1, 2, \dots, n_e\}$. The travel time between any two stops in the network, defined as the time elapsed between the departures of the same bus at those two stations, is modeled as a random variable. Let T_{e_i} be the random variable denoting the travel time on an edge $e_i \in \mathcal{E}$. Let $\mathcal{R}$ denote the set of predetermined routes in the transit network $\mathcal{G}$ where any $r \in \mathcal{R}$ is represented by a sequence of edges $r = (e_1^r, \dots, e_{n_r}^r), e_i^r \in \mathcal{E}$. Let $\mathcal{A}$ denote the set of all active trips in the network at any given instant of time.

Given any pair of stops $u, v \in \mathcal{V}$, a path from u to v is defined as a sequence of vertices and edges connecting the stop u to the stop v. Let $\mathcal{P}^{uv}$ denote the set of all such paths from stop $u \in \mathcal{V}$ to stop $v \in \mathcal{V}$. We are interested in providing the optimal routing choice for a passenger interested in traveling from an origin stop $o \in \mathcal{V}$ to a destination stop $d \in \mathcal{V}$, within a time budget of τ_b. While some of the paths in the set $\mathcal{P}^{od}$ could directly take a passenger from origin to destination on a single route, others might include a journey distributed on multiple routes, requiring the passenger to transfer from one bus to another. Let $\mathcal{F} \subseteq \mathcal{V}$ denote the set of stops with two or more different routes passing through it, hereafter referred to as the *transfer stops*. We note that while our definition requires a stop to be on more than one route to be considered as a transfer stop, in practice we could always extend the scope of the definition to include stops on different routes that are within predetermined walking distance of one another by introducing a pseudo-route with a deterministic travel time connecting the two stops. Further, given any two routes with multiple consecutive common stops, only the last stop in the sequence is considered to be transfer stop. Every trip $a^r \in \mathcal{A}$ has a scheduled time of arrival at each stop on its route. While the trip might not adhere to the schedule at every stop, we assume that in the case of early arrival at one of the transfer stops, the bus always leaves the transfer stops at the scheduled time by halting for the required amount of time. Further, we also assume that all the trips in the network operate with enough headway between trips to prevent the propagation of delays between consecutive trips of a bus.

With the recent growth of intelligent transportation systems, an increasingly large number of transit agencies are investing in Computer Aided Dispatch/Automatic Vehicle Locator (CAD/AVL) systems for their fleets. We assume that all the trips operating on every route in the network share their real-time departure times at stops with the transit command center. We also assume that we have access to the route configuration and static schedule data of the transit agency, often published in the General Transit Feed Specification (GTFS) format [15]. When a passenger requests a route to a destination while specifying a time budget, our proposed method retrieves the relevant real-time data from the command center along with the static schedule. It then finds the set of feasible paths for the trip and uses the proposed stochastic model to find the optimal policy generated by solving the multiobjective optimization problem. In the next section, we detail the stochastic model used for modeling the distribution of travel times and the steps involved in inferring these distributions from historical and real-time data.

III. Network Model

In this section, we model the travel time on the edges in the network using a probabilistic model to consider the variation of travel time on each edge. In particular, we consider the joint distribution of travel time on edges in the network to account for the spatial correlation of the travel time. Building on the intuition that any delay on a road segment due to traffic propagates to other connected road segments, we expect the travel times between different edges on a route to be positively correlated. The data collected in Champaign-Urbana validates our expectations. Fig. 1 shows an example of correlations of travel times between several edges on one of the routes. In the case of edges on two non-intersecting routes, we assume that their travel times are uncorrelated and exploring potential correlations to improve our network model is a subject of future work. Following the standard approach in the literature, we assume that the travel time on all the edges in the graph is a multivariate normal distribution [13]. The data from our case study, partly illustrated by Fig. 2, supports making such an assumption.

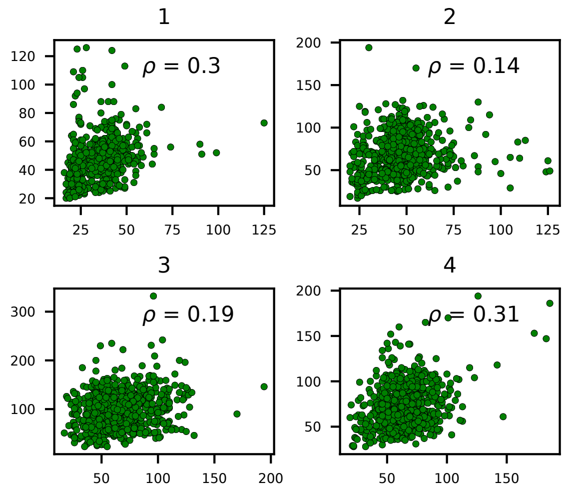

Fig. 1. Correlation of the travel time between several edges on the 'Green' route of CUMTD. Edges: **1** – (Green & Birch, Green & Orchard) and (Green & Orchard, Green & Busey); **2** – (Green & Orchard, Green & Busey) and (Green & Busey, Green & Gregory); **3** – (Green & Busey, Green & Gregory) and (Green & Gregory, Green & Goodwin); **4** – (Green & Second, Green & Locust) and (Green & Locust, Green & Neil).

Formally, let $\boldsymbol{T} = (T_{e_1}, \dots, T_{e_{n_e}})^T$ be a random vector representing the travel time on all the edges comprising the transit network $\mathcal{G}$ following a multivariate normal distribution. It follows that the travel time on every edge, T_{e_i}, by definition, is normally distributed. Let $\boldsymbol{T_r} = (T_{e_i^r}, \dots, T_{e_j^r})^T$ be a random vector representing the travel time on all edges in the route $r \in \mathcal{R}$ with $\boldsymbol{\mu_r} = \mathbb{E}[\boldsymbol{T_r}] = (\mu_{e_i^r}, \dots, \mu_{e_j^r})^T = (\mathbb{E}[T_{e_i^r}], \dots, \mathbb{E}[T_{e_j^r}])^T$ denoting its n_r-dimensional mean vector and $\boldsymbol{\Sigma_r} : \Sigma_r(e_i^r, e_j^r) = \mathbb{E}[(T_{e_i^r} - u_{e_i^r})(T_{e_j^r} - u_{e_j^r})]$ denoting its $n_r \times n_r$ dimensional covariance matrix. The mean vector $\boldsymbol{\mu_r}$ and the covariance matrix $\boldsymbol{\Sigma_r}$ that characterize the distribution of travel time on the route can be estimated using the historical travel time data samples gathered from AVL systems onboard the buses serving the route. Let m_r denote the number of such samples on a route r, each of the form $\boldsymbol{t_r^i} = (t^i_{e_1^r}, \dots, t^i_{e^r_{n_r}})^T, i \in \{1, ..., m_r\}$. Given these samples, the maximum likelihood estimates for the parameters of the multivariate normal distribution $\boldsymbol{T_r} \sim \mathcal{N}(\boldsymbol{\mu_r}, \boldsymbol{\Sigma_r})$ can be calculated using the following expressions [16]

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{r}} = \frac{1}{m_r} \sum_{i=1}^{m_r} \boldsymbol{t_r^i}, \tag{1}$$

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} = \frac{1}{m_r} \sum_{i=1}^{m_r} (\boldsymbol{t_r^i} - \hat{\boldsymbol{\mu}}_{\boldsymbol{r}})(\boldsymbol{t_r^i} - \hat{\boldsymbol{\mu}}_{\boldsymbol{r}})^T. \tag{2}$$

We note that equations (1) and (2) can be transformed into the following approximate recursive update rules for online parameter re-estimation eliminating the need for storing historical travel time data. Given a new route travel time sample $\boldsymbol{t'_r}$ and the number of samples m_r, we have

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{r}} \leftarrow \frac{m_r \hat{\boldsymbol{\mu}}_{\boldsymbol{r}} + \boldsymbol{t'_r}}{m_r + 1},$$

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} \leftarrow \frac{m_r \hat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} + (\boldsymbol{t'_r} - \hat{\boldsymbol{\mu}}_{\boldsymbol{r}})(\boldsymbol{t'_r} - \hat{\boldsymbol{\mu}}_{\boldsymbol{r}})^T}{m_r + 1}.$$

As a trip advances to its terminal stop, we observe the actual travel time on the traversed edges with the help of real-time data from the AVL system. The observed travel times,

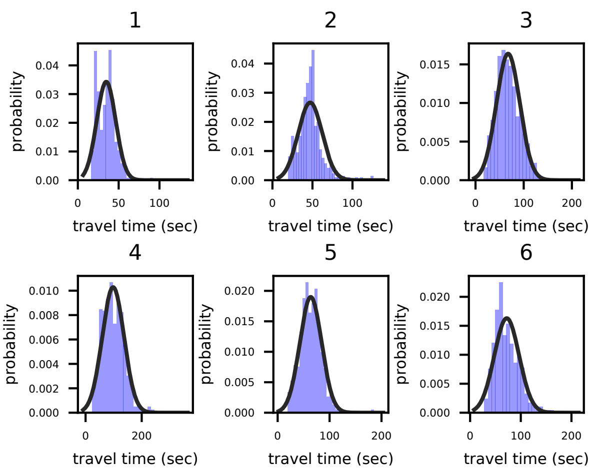

Fig. 2. Histogram of travel times on several edges in the 'Green' route of CUMTD fitted with a normal distribution. The edges are **1** – (Green & Birch, Green & Orchard); **2** – (Green & Orchard, Green & Busey); **3** – (Green & Busey, Green & Gregory); **4** – (Green & Gregory, Green & Goodwin); **5** – (Green & Second, Green & Locust); **6** – (Green & Locust, Green & Neil).

which are a consequence of the congestion levels on those edges, provide us with indirect insights into the currently prevalent traffic conditions on the entire route. Using these values of realized travel times on certain edges of a route, one can find the distribution of travel time on the rest of the edges conditioned on the observed data. In the case of a joint normal distribution, we know that the distribution of any subset of random variables conditioned on the realized values of the rest is also jointly normal [16]. Formally, consider two subvectors $\boldsymbol{T_r^a}$ and $\boldsymbol{T_r^b}$ of the random vector $\boldsymbol{T_r}$ where $\boldsymbol{T_r^a}$ consists of edges with recently observed travel times and $\boldsymbol{T_r^b}$ consists of edges whose travel times are yet to be observed. Further, the mean vector and the covariance matrix can be divided into the corresponding components,

$$\boldsymbol{\mu_r} = \begin{bmatrix} \boldsymbol{\mu_r^a} \\ \boldsymbol{\mu_r^b} \end{bmatrix}, \quad \boldsymbol{\Sigma_r} = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}.$$

Now, let $\boldsymbol{t_r^a}$ be the realized value of the random vector $\boldsymbol{T_r^a}$. Given this, the expression for mean vector and covariance matrix of the conditional distribution corresponding to the travel time on the unobserved edges $f(\boldsymbol{T_r^b} | \boldsymbol{T_r^a} = \boldsymbol{t_r^a}) = \mathcal{N}(\overline{\boldsymbol{\mu_r}}, \overline{\boldsymbol{\Sigma_r}})$, which follows from the fact that the conditional distribution is also jointly normal, can be given as [16]:

$$\overline{\boldsymbol{\mu_r}} = \boldsymbol{\mu_r^a} + \Sigma_{ab} {\Sigma_{bb}}^{-1} (\boldsymbol{t_r^a} - \boldsymbol{\mu_r^b}),$$
$$\overline{\boldsymbol{\Sigma_r}} = \Sigma_{aa} - \Sigma_{ab} {\Sigma_{bb}}^{-1} \Sigma_{ba}.$$

Given the conditional distribution of travel times on untraversed edges $\mathcal{N}(\overline{\boldsymbol{\mu_r}}, \overline{\boldsymbol{\Sigma_r}})$, we can obtain the reliability and expected travel time of any given path that better reflects the prevalent conditions. In the next section, we discuss our algorithm for online routing that evaluates the reliability and expected travel time of all feasible paths conditioned to the available data and then decides the optimal choice taking into account the passenger's preferences.

IV. Multiobjective Optimal Routing

In this section, we detail our approach for handling the expected travel time and reliability trade-off and present our algorithm for providing the optimal policy for routing given the passenger's delay tolerance. We use the weighted sum approach which transforms multiple objectives into an aggregated objective function by multiplying each objective function by a weighting factor and summing up all weighted objective functions [17]. The delay tolerance α can be used as the weighting factor to handle the trade-off between the two objectives. Consider a path p; let T_p denote the random variable representing the travel time on the path. Given the path travel time distribution, calculating the expected travel time, $\mathbb{E}(p) = \mathbb{E}(T_p)$, is trivial. The reliability, defined as the probability of reaching the destination within the time budget τ_b, can be given as $Re(p, \tau_b) = P(T_p \leq \tau_b) = F_{T_p}(\tau_b)$, where $F_{T_p}(t)$ denotes the cumulative distribution function of the random variable T_p. We now proceed to formalize the problem of multiobjective optimization for optimal routing.

Given an origin o, a destination d, a travel time budget τ_b, the distribution of travel time vector $\boldsymbol{T}$, the passenger's delay tolerance α, and a set of feasible paths $\mathcal{P}^{od}$, the objective function of interest can be formulated as follows:

$$\Theta(p, \tau_b) = \min_{p \in \mathcal{P}^{od}} \left[\alpha \mathbb{E}(p) \beta - (1-\alpha) Re(p, \tau_b) \right], \quad (3)$$

where $\mathbb{E}(p)$ and $Re(p, \tau_b)$ denote the expected travel time and the reliability of a path p given a time budget τ_b. The normalization factor β ensures that the magnitude of the expected travel time is comparable to that of the reliability. We note that when the passenger desires high reliability, i.e., has the delay tolerance $\alpha = 0$, the above formulation translates to the problem of maximizing just the reliability. Analogously, when the passenger's delay tolerance is $\alpha = 1$, problem (5) produces a standard LET route.

For any origin-destination pair in the network, the set of possible paths $\mathcal{P}^{od}$ often contains paths that involve one or more transfer stops $f \in \mathcal{F}$. We define a direct path as any path from origin to destination on a single route $r \in \mathcal{R}$ without any transfers. We first discuss the methods for evaluating expected travel time and the reliability of direct paths and then proceed to paths that involve transfers. Let $p = (e_i^r, \ldots, e_j^r)$ be a direct path from stop $o \in \mathcal{V}$ to another stop $d \in \mathcal{V}$. The path travel time T_p can be expressed as the sum of the jointly normal random variables representing the travel time on its constituent edges, $T_p = T_{e_i^r} + \ldots + T_{e_j^r}$. Consequently, the path travel time is also normally distributed, $T_p \sim \mathcal{N}(\mu_p, \sigma_p^2)$, with mean $\mu_p = \sum_{k=i}^{j} \mu_{e_k^r}$ and variance $\sigma_p^2 = \sum_{k=i}^{j} \sum_{l=i}^{j} \mathbf{\Sigma}(e_k^r, e_l^r)$. If a passenger is able to enter the bus immediately upon the passenger's arrival at the origin stop o, the expected travel time of the journey would be the same as the expected travel time of the path, $\mathbb{E}(p) = \mu_p$, and the reliability can be given as

$$Re(p, \tau_b) = P(T_p \leq \tau_b) = \Phi\left(\frac{\tau_b - \mu_p}{\sigma_p^2} \right),$$

where $\Phi(t)$ denotes the cumulative distribution function of the standard normal variable.

Recall that the above expressions for expected travel time and reliability of a path are based on the assumption that the passenger gets to depart immediately after arrival at the origin stop. In practice, however, this assumption rarely holds and the passenger is expected to wait until the arrival of the next bus on the route to begin the journey. We generalize the expressions for expected travel time and reliability of a path to account for the travel time of the bus from its last reported location to the origin stop. Consider an active trip $a^r \in \mathcal{A}$; let $w \in \mathcal{V}$ denote the last reported stop of the bus. Let $p' \in \mathcal{P}^{wd}, p' = (e_i^r, \ldots, e_j^r)$, be a direct path from the last reported stop of the bus to the destination and let τ_{elap} denote time elapsed since departure from stop w. The expressions for the expected travel time and reliability of the path can then be given as

$$E(T_{p'}) = \sum_{k=i}^{j} \mathbb{E}(T_{e_k^r}) - \tau_{elap} = \sum_{k=i}^{j} \mu_{e_k^r} - \tau_{elap},$$

$$Re(p', \tau_b) = \Phi\left(\frac{\tau_b - \mu_{p'}}{\sigma_{p'}^2} \right) - \Phi\left(\frac{\tau_{elap} - \mu_{p'}}{\sigma_{p'}^2} \right).$$

Unlike the direct path, the problem of evaluating the expected travel time and reliability on paths involving one or more transfers is challenging due to the inherent uncertainty in the travel time of all the buses involved in the transfer. In the remainder of this work, we only consider those paths that are either direct paths between origin and destination or involve a single transfer; we largely leave the case of multiple transfers for future work and briefly discuss it in Section VI.

We now build upon the generalized direct path case to derive the expressions for evaluating the expected time and the reliability of a single transfer path. Let $p \in \mathcal{P}^{od}, p = (p_{r_1}, p_{r_2}) = ((e_i^{r_1}, \ldots, e_j^{r_1}), (e_k^{r_2}, \ldots, e_l^{r_2}))$, denote a path over two routes $r_1, r_2 \in \mathcal{R}$ with a transfer at the stop $f \in \mathcal{F}$ that marks the end of edge $e_j^{r_1}$ and beginning of the edge $e_k^{r_2}$. Let τ_s be the start time of the journey and let $\gamma(a^{r_2}, f)$ denote the scheduled departure time of a trip $a^{r_2} \in \mathcal{A}$ at the transfer stop f. Let $w_{a^{r_2}} \in \mathcal{V}$ and $\tau_{elap}^{a^{r_2}}$ denote the last reported stop of the trip a^{r_2} and the time elapsed since departure at that stop. Let $p' \in \mathcal{P}^{w_{a^{r_2}} f}, p' = (e_m^{r_2}, \ldots, e_k^{r_2})$, be the direct path from $w_{a^{r_2}}$ to the transfer stop f. Similarly, let $w_{a^{r_1}} \in \mathcal{V}$ and $\tau_{elap}^{a^{r_1}}$ denote the last reported stop of the trip a^{r_1} and the time elapsed since departure at that stop. Let $p'' \in \mathcal{P}^{w_{a^{r_1}} f}, p'' = (e_n^{r_1}, \ldots, e_j^{r_1})$, be the direct path from $w_{a^{r_1}}$ to the transfer stop f. Let $\lambda = \gamma(a^{r_2}, f) - \tau_s$. Then, the expressions for the expected travel time and the reliability can be given as

$$E(p) = \max\left(\lambda, \sum_{z=m}^{k} \mu_{e_z^r} - \tau_{elap}^{a^{r_2}} \right) + \sum_{z=k}^{l} \mu_{e_z^r}, \quad (4)$$

$$Re(p, \tau_b) = Re(p'', \lambda) \Phi\left(\frac{\tau_b - \lambda - \mu_{p_{r_2}}}{\sigma_{r_2}^2} \right). \quad (5)$$

Algorithm 1 summarizes the above mentioned formulations for evaluating the expected travel time and reliability of any valid path of interest using edge travel time distributions conditioned on the observed travel time data.

Algorithm 1 *ObjEval*: Objective Function Evaluation

```
Input: A path p, time budget τ_b, observed travel times t_{a^r}, last reported stop(s) w, elapsed time(s) t_elap, random vector T, delay tolerance α
Output: Objective function value of the path p, Θ(p)
Let a^r ∈ A denote the trip under consideration
Let γ(a^r, s) denote the scheduled departure time of trip a on route r from the stop s
Let p ∈ P^{sd} denote the path under consideration
if p = (e_i^r, ..., e_j^r) then
    overline(T_r) ~ f(T_r | t_{a^r})
    E(p) = μ_p = Σ_{k=i}^{j} overline(μ_{e_k^r})
    σ^2(p) = Σ_{k=i}^{j} Σ_{l=i}^{j} Σ(e_k^r, e_l^r)
    Re(p, τ'_b) = Φ((τ_b − μ_p)/σ_p^2) − Φ((τ_elap − μ_p)/σ_p^2)
    Θ(p) = α E(p) β − (1 − α) Re(p, τ'_b)
end if
if p = (p_{r1}, p_{r2}) = ((e_i^{r1}, ..., e_j^{r1}), (e_k^{r2}, ..., e_l^{r2})) then
    overline(T_{r1}) ~ f(T_{r1} | t_{a^{r1}})
    overline(T_{r2}) ~ f(T_{r2} | t_{a^{r2}})
```

Evaluate $\mathbb{E}(p)$ using (4)
Evaluate $Re(p, \tau_b')$ using (5)
$\Theta(p) = \alpha\mathbb{E}(p)\beta - (1-\alpha)Re(p, \tau_b')$
end if
Return $\Theta(p)$

Given the procedure for evaluating the objective function value for any valid path, we now describe the algorithm for online adaptive routing in the transit network, presented as Algorithm 2. As stated earlier, we assume access to the departure time of the bus from every stop on its route using an onboard AVL system. Let τ denote the current time in the service day of the transit agency represented by the time interval $[\tau_{init}, \tau_{end}]$. Consider a passenger starting at origin stop $o \in \mathcal{V}$ at time τ_s intending to reach destination stop $d \in \mathcal{V}$ with-in the time budget τ_b. Let $s \in \mathcal{V}$ denote the latest reported stop of the passenger. The output of Algorithm 2 is an online adaptive policy that evaluates all feasible options at every transfer stop in an effort to find the path that best suits the passenger's requirements. In the next section, we implement our algorithm to provide optimal routing policy in the simulation of a transit network with real-world data.

V. Implementation and Case Study

In this section, we discuss a proof-of-concept implementation of the routing algorithm for the transit network in Champaign-Urbana. The Champaign-Urbana Mass Transit District (CUMTD), with its fleet of 111 buses, serves as the primary transit provider for the local community and the University of Illinois. The computer-aided dispatch/automatic vehicle location (CAD/AVL) units installed in the buses provide access to real-time arrival and departure information. CUMTD operates on 21 primary routes following a fixed schedule which includes appropriate layover time at the final stops of the routes to account for delays [18]. We use historical arrival/departure time data obtained from CUMTD and their published GTFS schedule [15] to model the travel time distributions and simulate bus operations in the transit network model. An implementation of PROTRIP and the simulator in Python along with the data used in modeling are available in a repository at github.com/pthangeda/protrip.

In this case study, we focus on three popular routes served by CUMTD, the Green route, the Orange route, and the Teal route. While both Green and Orange routes connect the downtown areas of Champaign and Urbana, the Green route runs through the parts of the city with high foot and bike traffic whereas the Orange route runs on an arterial road with less pedestrian interference. Further, the section of the Green route from the transfer stop f to destination d faces higher pedestrian interference and the resulting delays when compared to the other section. On the other hand, the Teal route faces high traffic until transfer stop f and then faces lower interference until it reaches the destination stop d. We use the GTFS schedule published on January 30, 2019 [18] to extract the information of the stops, the routes operated by CUMTD, and the scheduled departure time of all trips from their origin stop and all the transfer stops along their routes.

Algorithm 2 Multiobjective Optimal Routing

Input: $o, d \in \mathcal{V}$, time budget τ_b, delay tolerance α
Output: Policy $\Pi(f, \tau), f \in \mathcal{F} \cup \{o\}, \tau \in [\tau_{init}, \tau_{end}]$
Let w_{a^r} be the last recorded stop of the trip $a^r \in \mathcal{A}$
Let η_{a^r} be the departure time at w_{a^r} of the trip $a^r \in \mathcal{A}$
Let $\boldsymbol{t}_{a^r}$ be the observed travel time of the trip $a^r \in \mathcal{A}$
Initialize location of the passenger $s \leftarrow o$
Initialize $\tau \leftarrow \tau_{init}$, $\tau_b' \leftarrow \tau$
repeat at every time instant τ
  for all active trips $a^r \in \mathcal{A}$ **do**
    Retrieve the latest reported stop $w' \in \mathcal{V}$
    if $w' = w_{a^r}$ **then**
      $\tau_{elap} \leftarrow \tau - \eta_{a^r}$
    end if
    if $w' \neq w_{a^r}$ **then**
      $\boldsymbol{t}_{\boldsymbol{a^r}}(w_{a^r}, w) \leftarrow \tau - \eta_{a^r}$
      $w_{a^r} \leftarrow w'$, $\eta_{a^r} \leftarrow \tau$
    end if
  end for
  if $s \in \mathcal{F} \cup \{o\}$ **then**
    for all $p \in \mathcal{P}^{sd}$ **do**
      $\Theta(p) \leftarrow ObjEval(p, \tau_b')$
    end for
    $\Pi(s, \tau) \leftarrow \min_p \Theta(p)$
    $\tau_b' \leftarrow \tau_b - (\tau - \tau_s)$
  end if
  for all active trips $a^r \in \mathcal{A}$ **do**
    if a^r at final stop of $r \in \mathcal{R}$ **then**
      $\hat{\boldsymbol{\mu}}_{\boldsymbol{r}} \leftarrow \frac{m_r \hat{\boldsymbol{\mu}}_{\boldsymbol{r}} + \boldsymbol{t}_{a^r}}{m_r + 1}$
      $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} \leftarrow \frac{m_r \hat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} + (\boldsymbol{t}_{a^r} - \hat{\boldsymbol{\mu}}_{\boldsymbol{r}})(\boldsymbol{t}_{a^r} - \hat{\boldsymbol{\mu}}_{\boldsymbol{r}})^T}{m_r + 1}$
      $\mathcal{A} \leftarrow \mathcal{A} \setminus \{a^r\}$
    end if
  end for
  return $\Pi(s, \tau)$
  $\tau \leftarrow \tau + \Delta\tau$
until end of service day $\tau = \tau_{end}$

Using this data, we build our graph representing the transit network in the city and establish the scheduled departure time of all trips at all the transfer points. To build the stochastic model of the travel times on these routes, we use historical departure data of 500 trips provided by CUMTD, amounting to an average of about 150 travel time data samples on each edge. Using the expressions for maximum likelihood estimation of parameters stated in (1) and (2) and the data samples, we infer the initial model of the travel time.

To demonstrate the multiobjective routing algorithm as presented in Algorithm 2, we analyze the journey of a passenger intending to travel from downtown Urbana (marked o in Fig. 3) to downtown Champaign (marked d in Fig. 3). For the sake of clarity, we consider the following three paths from the stop o to the stop d: 1) the direct path on Green route, 2) the direct path on Orange route, 3) the path starting with Green route and transferring to the Teal route at stop f. In order to simulate real-world bus

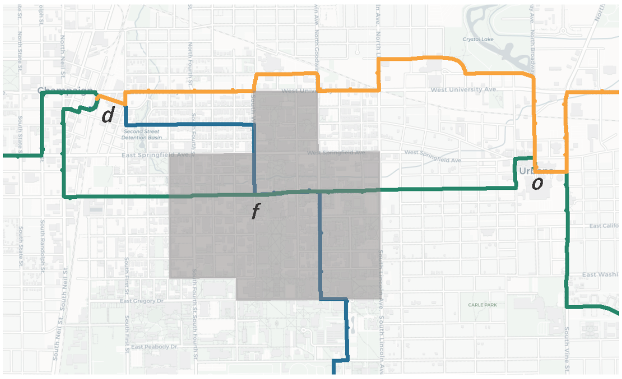

Fig. 3. Visualization of the three routes considered in the case study: Green, Orange, and Teal. Letter o denotes the origin stop located in downtown Urbana, f denotes a transfer stop, and d denotes the destination stop located in downtown Champaign. The highlighted area indicates the parts of city with irregular and high pedestrian traffic.

movement in this transit network, for every route $r \in \mathcal{R}$, we sample from the multivariate distribution representing the travel time distribution $\boldsymbol{T_r}$ and then use these samples as the observed travel time for trips departing the origin stop of the route according to the static schedule. We assume that the passenger arrived at the origin stop at 6:00 PM and desires to reach the destination in 10 minutes, i.e., 600 seconds. Under the simulated conditions, the routing algorithm selects the direct path on the Green route when the passenger has a high tolerance to delays ($\alpha = 1$) and selects the direct path on the Orange route when the passenger has no delay tolerance ($\alpha = 0$). Given a moderate delay tolerance ($\alpha = 0.4$), the algorithm selects the path with a transfer at f. These results are in line with the observations in the actual system where high pedestrian traffic on the Green route from the stop f to the stop d increases the estimated travel time of that segment of the route.

VI. Conclusion and Future Work

We presented an algorithm that generates an adaptive routing policy for a passenger constrained by a travel time budget and with a known tolerance to potential delays. We also used a realistic model for the travel times in the transit network, accounting for the correlation of travel time between different edges on a route. The proposed framework provides a routing policy that suits the preferences of the passenger, as demonstrated using a proof-of-concept implementation of the algorithm with real-world data.

While this work develops the framework for multiobjective optimal routing incorporating passenger's delay tolerance, still much work remains to be done for deploying the tool in massive real-world transit networks. Computational efficiency is crucial while scaling to serve potentially thousands of simultaneous requests and we plan to focus on designing preprocessing techniques and improving subroutines' efficiency to significantly reduce the overall computation time. Lack of consideration of paths that contain more than one transfer stops is one of the main current shortcomings of our algorithm. While we could potentially extend the current algorithm to cover paths with multiple transfers by considering, among other things, the availability of near-optimal alternate routes in the case of a missed transfer, we leave the details of this important extension to future research. Finally, providing directions from any geographical location to any other location is one of the key requirements to increase the acceptability of our tool. We plan to include a user-specific walking speed distribution, learned over time from the data provided by the user, to provide an end-to-end risk-aware optimal routing solution for urban travelers.

Acknowledgment

The authors would like to thank Jay Rank and Evan Alvarez at CUMTD for providing the data and insights of their operations.

References

[1] R. J. Javid and R. J. Javid, "A framework for travel time variability analysis using urban traffic incident data," *IATSS Research*, vol. 42, no. 1, pp. 30–38, 2018.

[2] S. Susilawati, M. A. Taylor, and S. V. Somenahalli, "Distributions of travel time variability on urban roads," *Journal of Advanced Transportation*, vol. 47, no. 8, pp. 720–736, 2013.

[3] R. W. Hall, "The fastest path through a network with random time-dependent travel times," *Transportation Science*, vol. 20, no. 3, pp. 182–188, 1986.

[4] N. Borole, D. Rout, N. Goel, P. Vedagiri, and T. V. Mathew, "Multi-modal public transit trip planner with real-time transit data," *Procedia-Social and Behavioral Sciences*, vol. 104, no. 775-784, p. 2nd, 2013.

[5] H. Frank, "Shortest paths in probabilistic graphs," *Operations Research*, vol. 17, no. 4, pp. 583–599, 1969.

[6] P. B. Mirchandani, "Shortest distance and reliability of probabilistic networks," *Computers & Operations Research*, vol. 3, no. 4, pp. 347–355, 1976.

[7] Y. M. Nie and X. Wu, "Shortest path problem considering on-time arrival probability," *Transportation Research Part B: Methodological*, vol. 43, no. 6, pp. 597–613, 2009.

[8] S. Samaranayake, S. Blandin, and A. Bayen, "A tractable class of algorithms for reliable routing in stochastic networks," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 199–217, 2012.

[9] Y. Nie and Y. Fan, "Arriving-on-time problem: discrete algorithm that ensures convergence," *Transportation Research Record*, vol. 1964, no. 1, pp. 193–200, 2006.

[10] C. Carrion and D. Levinson, "Value of travel time reliability: A review of current evidence," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 4, pp. 720–741, 2012.

[11] M. Ramezani and N. Geroliminis, "On the estimation of arterial route travel time distribution with Markov chains," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1576–1590, 2012.

[12] B. J. Gajewski and L. R. Rilett, "Estimating link travel time correlation: an application of Bayesian smoothing splines," *Journal of Transportation and Statistics*, vol. 7, no. 2-3, pp. 53–70, 2005.

[13] W. Dong, H. L. Vu, Y. Nazarathy, B. Q. Vo, M. Li, and S. P. Hoogendoorn, "Shortest paths in stochastic time-dependent networks with link travel time correlation," *Transportation Research Record*, vol. 2338, no. 1, pp. 58–66, 2013.

[14] Y. Fan, R. Kalaba, and J. Moore II, "Shortest paths in stochastic networks with correlated link costs," *Computers & Mathematics with Applications*, vol. 49, no. 9-10, pp. 1549–1564, 2005.

[15] Google Inc., "General Transit Feed Specification Reference," 2012, [Accessed 12-October-2019]. [Online]. Available: https://developers.google.com/transit/gtfs/reference

[16] Y. L. Tong, *The Multivariate Normal Distribution*. Springer Science & Business Media, 1990.

[17] K. Miettinen, *Nonlinear Multiobjective Optimization*. Springer Science & Business Media, 2012, vol. 12.

[18] "CUMTD open data website," [Accessed 20-October-2019]. [Online]. Available: https://developer.cumtd.com