



Center for Advanced Multimodal Mobility Solutions and Education

UTC Project Information – CAMMSE @ UNC Charlotte	
Project Title	Using Computational Biology to Mitigate Path Overlap in Transit Assignment
University	The University of Connecticut
Principal Investigator	Nicholas Lownes and Timothy Becker
PI Contact Information	(860)-486-2717 / nicholas.lownes@uconn.edu and (860)768-4502 / tbecker@hartford.edu
Funding Sources and Amount Provided (by each agency or organization)	The University of North Carolina at Charlotte: \$44,958 The University of Connecticut: \$24,847
Total Project Cost	\$69,805
Agency ID or Contract Number	
Start and End Dates	10/01/2019 – 09/30/2021
Brief Description of Research Project	In the past thirty years, transit assignment has grown from a niche theoretical exercise to a rigorous field of study with a variety of advanced models and methodologies designed to incorporate the unique decision-making process of transit trips in the planning process. A significant boon to the transit modeling community came with the advent of the General Transit Feed Specification (GTFS) in 2006. Prior to 2006, information on transit networks was difficult to obtain and was inconsistently formatted – meaning that software developed for one network may need significant modifications to be applied elsewhere. The motivation of having



Center for Advanced Multimodal Mobility Solutions and Education

transit systems searchable on Google Maps proved great enough to convince hundreds of transit systems around the globe to develop their own GTFS feed and make it openly available. This opened the door to a vast amount of research potential, not least the transit assignment community. One particular software, FAST-TriPs, leveraged the structure of GTFS to create a dynamic transit passenger assignment (DTPA) program with improved computational performance than previous tools. The program takes a detailed (or enhanced) version of GTFS and transit demand data and simulates the transit choices people might make by generating many possible paths for every trip a simulated traveler wants to make and then assigns them to paths probabilistically, incorporating the capacity of the transit vehicles along the way.

In this path generation phase, one issue identified by researchers working to take DTPA from research to practice was the presence of path overlap. Path overlap occurs when two or more paths in a path choice set contain many similar features with slight variations, violating independence assumptions in the logit-based path choice model and potentially resulting in a poor representation of transit path choice behavior. In the path generation phase there are also sometimes nonsensical paths included in the path set (multiple transfers between two routes; alighting and then reboarding the same route) that waste computational resources and in some cases (albeit rarely) result in one of those trips being chosen in the



Center for Advanced Multimodal Mobility Solutions and Education

DTPA.

GTFS organizes the transit network as a collection of linked spreadsheets that organizes the system as stops, trips and routes. Trips are a sequence of stops visited by a transit vehicle at a particular time of the day. It is these trips, and their associated stop times that are used by FAST-TrIPs in the path generation phase. Because GTFS presents these trips as sequences, it presents an opportunity to leverage work with sequences to potentially improve on the path generation phase of DTPA.

Computational molecular biology and computer science are fields that have been working to efficiently operate with strings and sequences for decades. Among others, the Longest Common Subsequence (LCS) algorithm is a method used to measure the similarity between a pair, or group of sequences. The sequences can be either sequences of the letters T, A, C and G representing the bases of a DNA molecule or a sequence of stops representing a potential transit trips. Incredibly, this dynamic-programming based method run in linear time – making them very computationally efficient.

The goals of the proposed research are broadly to apply the methods of sequence alignment and matching from computational biology, focusing on LCS, to the DTPA problem in three contexts:



Center for Advanced Multimodal Mobility Solutions and Education

	<ol style="list-style-type: none"> 1. A priori characterization of the network 2. Elimination of nonsensical paths during path generation phase 3. Mitigation of path overlap through development of metrics to be integrated with the path choice model.
<p><i>Describe Implementation of Research Outcomes (or why not implemented)</i></p> <p><i>Place Any Photos Here</i></p>	
<p><i>Impacts/Benefits of Implementation (actual, not anticipated)</i></p>	
<p><i>Web Links</i></p> <ul style="list-style-type: none"> • <i>Reports</i> • <i>Project website</i> 	<p>https://cammse.uncc.edu/sites/cammse.uncc.edu/files/media/CAMMSE-UNCC-2020-UTC-Project-Information-10-Lownes-Becker.pdf</p> <p>https://cammse.uncc.edu/sites/cammse.uncc.edu/files/media/CAMMSE-UNCC-2020-UTC-Project-Report-10-Lownes-Becker-Final.pdf</p>