Psychology of Route Choice in Familiar Networks: Minimizing Turns and Embracing Signals

Mohan Venigalla, Ph.D., P.E., M.ASCE[1]; Xi Zhou, Ph.D., P.E., M.ASCE[2]; and Shanjiang Zhu, Ph.D., M.ASCE[3]

Abstract: It is widely accepted that path choice of a trip is dependent on trip characteristics, network attributes, and a traveler's personal characteristics. The best-known network variables that influence route choice are travel distance and travel time. This research attempts to study the influence of other network variables, namely signals, turns, and roadway classification on route choice. Real-world trip data from path trajectories tracked by a global positioning system (GPS) in an urban area are used to isolate nearly 5,700 unique real paths. Procedures to compute the theoretical shortest time path (STP) and shortest distance path (SDP) based on travel time and distance as impedance variables, respectively, are developed. Street network data are augmented with data on signalized intersections. Procedures to identify turns and road classes along the real and theoretical paths, and methods to quantify turn penalties are developed. The real paths are compared to their STP and SDP counterparts to identify discernible relationships between the network variables and the path choice. The number of traffic signals along the path is found not to be a statistically significant factor during the path selection process. In contrast, for trips shorter than 16.1 km (10 mi), drivers embraced more signals along their chosen travel paths than the number of signals that are present along the SDP or STP. It is observed that drivers are willing to spend longer time or travel longer distances on paths that have fewer turning movements. Furthermore, there is statistical evidence to indicate that real paths have fewer turns per 1.61 km (1 mi) than both STP and SDP. When they must make a turn to complete their trips, drivers seem more prone to making the turn at a signal-controlled intersection, while at the same time trying to minimize the number of turns occurring at nonsignalized intersections. Most notably, for trips shorter than 8.05 km (5 mi) in length, real paths have a statistically significant fewer left turns per 1.61 km (1 mi) than right turns per 1.61 km (1 mi). This leads to the conclusion that drivers tend to minimize left turns while selecting a path. Exceptions to these observations are very few and they all happen with paths longer than 16.1 km (10 mi). It is anticipated that the findings from this research will be influential and make it easier to find paths that are more consistent with drivers' real choices and consequently provide more sound solutions to modeling transportation networks. **DOI: 10.1061/(ASCE)UP.1943-5444.0000364.**

Author keywords: Route choice; Transportation networks; Influence of signals on route choice; Influence of turns on route choice; Influence of road class on route choice; Path-finding algorithms; Global positioning system (GPS)-tracked trip trajectories.

Introduction

On a typical street network, travelers are faced with a choice among multiple paths for travel between the origin and destination (O-D pair) of their trip. Given the characteristics of the trip (purpose, time, origin, destination, travel mode, etc.); attributes of the alternative routes (network variables); and a traveler's personal characteristics, the trip maker chooses the "best" route through the transportation network following some criteria (Antonisse et al. 1989). This route-choice problem may also be stated in the form of a rhetorical question, "What is the most practical path from point A to point B?" However, it is widely accepted that there is no single answer to this question. The best-known network variables that influence route choice are travel distance and travel time. Route-choice criteria in the real world are not necessarily based on the path with the shortest travel time, travel cost, or a combination thereof (impedance). Rather, most drivers choose a route that they perceive to be the best according to their personal preferences, knowledge, and experience (Liu 1996).

The route-choice behavior is of interest to transportation modelers. The traffic assignment methods used in the traditional four-step transportation planning process require an adherence to certain path or route-choice behavior. The framework proposed by Wardrop and Whitehead (1952) provided a theoretical basis for formulating and solving the traffic assignment problem as user equilibrium (UE) and system optimal (SO) optimization problems (Chatterjee and Venigalla 2003). Wardrop's first and second principles of equilibrium, popularly known as the Wardrop criteria, assume a generalized route-choice behavior of individuals (Venigalla et al. 1999).

Both UE and SO problems involve only the minimization of individual or system-wide impedance. Despite the popularity of the UE method in present-day travel demand modeling exercises, there is no empirical evidence that validates real-world route-choice behavior conforming to the Wardrop criteria (Correa and Stier-Moses 2011). Research has shown that numerous criteria, which could be used to formulate a route and route-choice application, are more stochastic than deterministic in nature. Therefore, assuming travel time (or distance) as the sole criterion of route choice may be an overly simplistic abstraction of individual driver behavior and may result in an inaccurate representation of traffic in transportation models.

A number of probabilistic route-choice models have also been developed and studied on the basis of the Wardrop criteria. A few

[1]Associate Professor, George Mason Univ., Fairfax, VA 22030 (corresponding author). E-mail: mvenigal@gmu.edu

[2]Transportation/Traffic Engineer, CH2M HILL, Inc., 15010 Conference Center Dr., Suite 200, Chantilly, VA 2015.

[3]Assistant Professor, George Mason Univ., Fairfax, VA 22030.

Note. This manuscript was submitted on October 8, 2015; approved on July 12, 2016; published online on August 26, 2016. Discussion period open until January 26, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Urban Planning and Development***, © ASCE, ISSN 0733-9488.

route-choice models are also derived from utility theory, which states that each person tries to maximize utility when faced with a choice among competing routes. On the basis of the relative importance of factors of influence, the route-choice model first identifies the set of sufficiently attractive alternatives for specific travelers. Travelers make their choices from this set, with the chosen route being the one that best satisfies their needs and is consistent with their personal constraints and preferences. A variety of logit/probit route-choice models were developed that vary by the basic structure of the model. These models include multinomial probit (MNP) and multinomial logit (MNL) models (Daganzo and Sheffi 1977); the C-logit model (Cascetta et al. 1996); the implicit availability/perception (IAP) logit model (Cascetta and Papola 1998); and the path-size logit model (Ben-Akiva and Bierlaire 1999). Other commonly used logit-based models include the PCL model (Chu 1989); the CNL model (Vovsha 1997); and the logit kernel or mixed logit model (McFadden and Train 2000; Ben-Akiva and Bolduc 1996).

Other than the best-known factors, travel distance and travel time, research has shown that route-choice behavior is also influenced by demographic variables (such as age, gender, profession, or household structure); road and traffic conditions (such as time of day, travel cost, road classification, or congestion); trip characteristics (such as time of day, purpose, or mode), and environment conditions (such as weather or accident) (Jan et al. 2000).

Studies have been conducted to identify the influence of travel-time reliability (Jackson and Jucker 1982), travel information (Stern et al. 1993; Abdel-Aty et al. 1994; Polydoropoulou et al. 1994; Srinivasan and Mahmassani 2000; Chen and Jovanis 2003), and usage of freeways (Li 2004) on path-choice behavior. Jackson and Jucker (1982) found that travel-time reliability, defined as the difference between the 90th percentile and the median travel times, could be an important influence factor on commute route choices. Travel-time reliability may be positively correlated with criteria like number of traffic signals along a route or the safety of the route. However, the study did not explore this relevance further.

The decision-making process in route choice is a learning process, which is central to the driver's cognition (Polydoropoulou et al. 1994). Therefore, information acquired through the experience of earlier travel choices is considered when a driver is making the next decision. Srinivasan and Mahmassani (2000) have discussed the influence of travel information on route choice theoretically. Abdel-Aty et al. (1994) conducted a stated preference survey in Los Angeles that indicated that about 60 percent of drivers listen to a travel information report. The study did not indicate how the drivers responded to this information. Two other surveys in Sweden and Israel found that on average two-thirds of commuters will change their travel behaviors based on real-time information (Stern et al. 1993).

Chen and Jovanis (2003) investigated drivers' responses to incident and congestion information by studying whether drivers would follow the guidance to change routes if they were advised to turn or switch to a freeway. Although the study found that the influence of advice on turns is significant, it did not reveal whether drivers tried to minimize the number of turns along the whole route, as the guidance suggests only which road link drivers should take immediately next without showing guidance for the rest of the route.

An investigation of route-choice behavior on morning commutes (Li 2004) compared routes that commuters most frequently chose over alternatives. Statistical analysis showed that the primary routes employed a higher freeway percentage and fewer signals than alternative routes. However, primary routes have not been compared to the computed shortest paths.

Jan et al. (2000) have used global positioning system (GPS) data to investigate variations in path choice. They found that travelers often took paths that greatly deviated from the shortest paths, but they did not explain why travelers choose these routes.

Compared to suburban or rural areas, urban areas tend to have a much higher density of intersections. Therefore, the number of traffic lights and turning movements along the path may significantly affect route choice. Zhuang et al. (2012) examined 50 trips between four O-D pairs, and found that the experienced routes (GPS tracking routes used by taxis) have lower frequencies of signalized intersections and turning movements than theoretical shortest-path routes. However, there is still a deficiency in the literature with respect to the effects of these two specific factors. Past studies did not reveal how these factors affected path choice or qualitatively identify these effects.

A number of studies based on stated preference surveys have been carried out to identify factors influencing route choice other than travel time and distance. Some studies pointed out that drivers tend to minimize signals and turns along the path. However, very few studies presented any empirical evidence based on field data to support this observation. Among those studies that considered empirical data, either the sample sizes were too small or these studies did not investigate influential factors in depth. Most recently, Zhou (2014) and Zhou and Venigalla (2014) studied the effect of traffic signals and turning movements along the path on route choice.

This research aims at analyzing a large data set of real-world trip data to study the influence of the aforementioned network variables on route choice. The study objective is to determine the extent to which turn movements, signalized intersections, and roadway type have influence on route choice. Furthermore, the study will explore the influence of roadway class on path choice. The scope of the study includes statistical analyses on real-world trips tracked by GPS equipment to examine the influence of signals and turning movements on drivers' path choices in an urban street network. Real-world trips made in the metropolitan area of Minneapolis-St. Paul (Twin Cities), Minnesota, during the period from September to December 2008 and the associated network data are extracted and processed with customized geographic information systems (GIS) applications.

Methodology

Fig. 1 illustrates the study methodology employed for examining the factors influencing drivers' real-world path choices. As the schematic indicates, given a digitized street network in the form of a set of nodes and links, and trip trajectory data (as a series of points) collected by GPS devices mounted on vehicles, the real paths that drivers actually chose can be identified. It is further hypothesized that, by conflating the street network attributes such as signalized intersections and road classes with the paths, additional insights into route-choice behavior can be gleaned.

GPS-tracked path trajectories collected for a study in the metropolitan area of Twin Cities (or study area) during the period from September to December 2008 were acquired. The data set contains information on trips made by 44 randomly selected volunteers in the 21–65 age group. All the drivers were local to Twin Cities and presumably have fairly good knowledge of the alternative routes in the network. The volunteers commuted alone and made travel choices without any instructions. A GPS device was installed in the vehicle of each study participant. The device recorded trajectories of each vehicle at a frequency of one GPS location point per

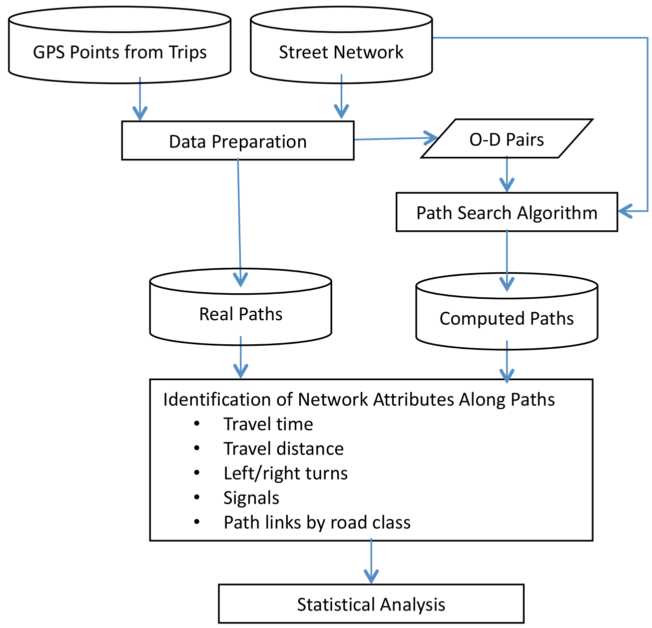

Fig. 1. Study methodology for analysis of factors influencing path choice

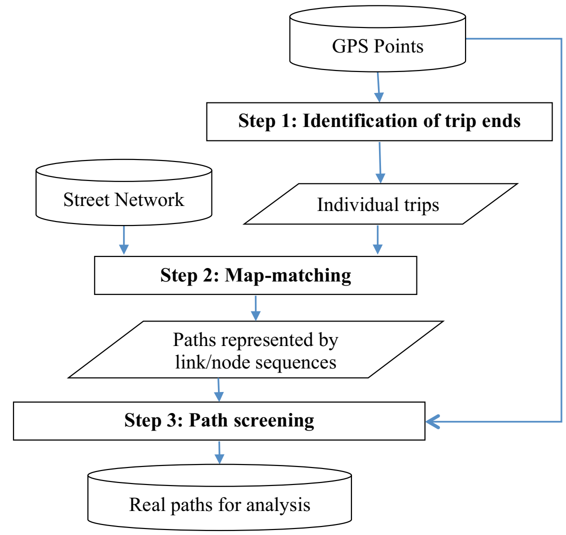

Fig. 2. Process to identify real paths from GPS-tracking data

second. The geographic location and time stamps of each point were documented and projected onto the *ArcGIS* shape files for postprocessing.

The street network data covering seven counties in the study area were obtained from the U.S. Geological Survey's (USGS) National Geospatial Program. The network contains full paths for all trips in the data set. The GPS points were first snapped to the street network, generating the set of real paths. Each real path was then represented by a sequence of nodes, as well as a sequence of links. The first and last nodes of each path were identified as the origin and destination of the trip, respectively. GPS data were analyzed to compute turn penalties for incorporation in path search algorithms. Using Python scripting, Dijkstra's algorithm (1959) was implemented as the path search algorithm, which was then used for finding the shortest path between each O-D pair in the data set. The GPS time-stamps were used to obtain travel speeds on road links, by hour of day. For links where GPS trajectory data were not available, posted speed limits were used to compute travel times. Thus computed travel times were used in path search algorithms where travel time was used as an impedance.

The data on signals were acquired separately. An algorithm to identify left turns, right turns, and through movements has been developed. By overlaying on the road network, network variables that are relevant for the paths in the solution space and turns were identified and tagged appropriately. Paired sample t-tests between real paths and shortest paths were then conducted on these factors to analyze their influence on path choice.

Real Path Identification

A three-step methodology was developed to identify real paths as shown in Fig. 2. The first step generates individual trips from the data set containing GPS-tracking points. Second, these trip points were snapped to the street network by a map-matching algorithm, yielding paths represented by both node sequence and link sequence. In the last step, further path screening eliminated invalid paths, producing the set of real paths for later analysis.

Identification of Trip Ends

The vehicle trajectories in the database were first distinguished trip by trip for subsequent analysis. Ideally, if the interval between two successive points was more than a few seconds, the two points would be treated as points belonging to two different trips. However, when a satellite signal was lost while the driver was still on the trip, the data on trip trajectories may have been broken. This could lead to erroneously splitting a single trip into more than two trips. Another exception could also occur. If the driver finished a trip when the GPS device was still on, the dwelling time would be included and the two trips would be treated as one.

Processing of trip data to identify individual trips from the large set of GPS data involved the following effort:

1. It was assumed that the vehicle may not stay on a digitized road at the end of a trip. This assumption is expected to help screen out the data pertaining to trips in which the GPS device was still on when the driver had completed a trip. As the trajectories were not aligned with a network link, a threshold was needed. With trial and error and a random manual check, it was determined that a threshold of 30 m could provide a good measure. Thus identified GPS points that were away from the road network were removed from the data set.
2. Determining the minimum possible time gap required an identification process for the next trip. Intuitively, there must be a threshold below which a new trip was not possible. However, if the gap was larger than the threshold, further analysis would be needed to determine whether it was because of a trip's end or just a period of signal loss. Previous research showed that 30 s was a good threshold for the minimum time gap (Du and Aultman-Hall 2007).
3. Distinguishing signal loss from a trip's end for the time gap longer than the minimum threshold of 30 s was the third important step in identifying individual trips. If the time gap was caused by signal loss, the average speed of the vehicle during this gap would not be much different from average speeds before and after the gap, if the driving pattern was assumed constant. The average speed during the gap can easily be estimated from the time and distance recorded by GPS devices. The highest free flow speed was 112 km/h (70 mi/h), and average speeds on most streets are no more than 80 km/h (50 mi/h). Therefore, it would be reasonable to identify a trip's end when the average speed during the time gap was 50% less than speeds before and after

the gap. Du and Aultman-Hall (2007) recommended using 20 points before and after the gap to calculate the assumed driving speed. This proved enough to obtain successful identification.

This methodology has proved to be very effective in identifying individual trips. However, certain special situations posed challenges to this methodology. For example, if the vehicle met with traffic congestion or traveled from an expressway to a local road when the GPS device was experiencing a signal loss, a continuous trip was split when the average speed reduction was over the threshold. Such instances were assumed to be rare and could not be avoided. A few random checks have shown that this identification process could yield better results than the results obtained only by using the original time stamps recorded by GPS devices.

Map-Matching Algorithm

A route was obtained by connecting GPS points according to the time order in which they are recorded. However, the route did not match any links on the street network in many cases, because of either an error in the GPS location or an inaccurate digital road network. It was necessary to snap the GPS points of each trip to the digital street network by the map-matching technique. Real trip paths were identified in the format of node and/or link sequences between trips' origins and destinations.

Numerous studies have developed procedures to perform map-matching effectively and accurately (Bernstein and Kornhauser 1996; Quddus et al. 2006, 2007; White et al. 2000). Based on the recommendations of these studies, embedded tools in ArcGIS were used extensively for map matching. Since the GPS collected points second by second, multiple points may have had a common nearest link; thus, the link would not be kept in the sequence repeatedly if it was the same as its previous link. Where necessary, custom tools within or outside the ArcGIS environment were developed using Python scripts.

The set of candidate links identified by ArcGIS may have contained links that were not on real paths used by trips in the data set. On the other hand, a few links forming a real path may have been missing, because neither the GPS device nor the digital street map was 100% accurate and/or compatible. To screen out the incorrect links and find the missing links, further processing of the digital network was needed. This processing required examination of a connection between successive links to make sure the link sequence was consistent with the real travel route and direction. The map-matching algorithm developed for this study (also in Python script) eliminated wrong links, while at the same time retrieving missing links. The algorithm also generated a new node sequence that was consistent with the actual trip.

About 8% of total trips (1,668 out of 20,174 trips) failed in map matching because of errors attributable to the digital street network (for example, topological errors, missing links, or incorrect link configuration).

Path Screening

The map-matching process reduced the data to 18,560 trips from 20,174 trips. For identifying traversable paths in the network, the data were further screened by adhering to the following criteria and by eliminating the trips that did not fit the criteria:

1. A path requires at least two links. If a path contained fewer than three nodes, the trip was eliminated from the analysis data set.
2. Trips with path lengths shorter than 1 min in travel time indicate a potentially faulty GPS device and therefore were dropped from the analysis data set.
3. If multiple paths made by the same driver were identical to each other, they were assumed to be commuting trips. Since such duplicate trips did not provide additional insights into the route-choice behavior of the driver making those trips, only one of these multiple trips was included in the analysis data set.

After this screening process was completed, the remaining 5,694 trips were identified with a valid path represented by both a link sequence and a node sequence. The first and last nodes in the path were tagged as the origin and the destination, respectively. This O-D node set was then used for computing theoretical shortest paths.

Fig. 3 depicts the distribution of street links that were used by one or more paths. This map indicates that identified valid paths primarily occurred in the urban areas. Links used over 100 times by the trips in the analysis data set (thick red lines) were mostly within downtown areas. In the suburban areas, the most used links were primary or secondary roads, and none of them was used more than 10 times.

Computation of Shortest Paths

Custom Python scripts were developed for implementation of shortest path algorithms. The script uses Forward Star notation to describe the street network. The notation represents the network as an adjacency list:

$$G = \{n \in N, n = \text{Node}[id, x, y, \text{signal}, FS(n)]\} \tag{1}$$

where G = graph; n = node that belongs to a set of nodes N, with an identification tag id, coordinates x and y, signal existence, and a link set named *forward star*.

The forward star of a node is the set of all links starting from this node. The forward star notation is a representation of network connectivity. The Dijkstra algorithm (Dijkstra 1959) was used in the script for finding shortest paths.

For a given O-D pair in the trip set derived from the real path identification (i.e., for each real trip in the data set), the script generated two paths. The first path was based on shortest travel time [shortest time path (STP)], and the second path was based on shortest distance [shortest distance path (SDP)]. As in the case of real paths, each of the two computed shortest paths was represented by a node sequence as well as by a link sequence. The link sequence or node sequence along these three paths (real path, STP and SDP) may or may not have been the same. Comparing the node/link sequences of any two paths can identify whether or not they are identical to each other. Two paths can be regarded as identical only when they have the same number of nodes/links in their sequences, and when the two paths have the same nodes/links at exactly the same points in those sequences.

Signals and Turns along Paths

The signal information was obtained from the local jurisdiction responsible for signal operations in the Twin Cities. The signal data, which were a point feature among the GIS layers used in the analysis, were conflated to the nearest intersection based on the coordinates. Each intersection node in the network was tagged with a Boolean attribute to indicate whether a signal existed at that location. Such tagging enabled the tracking of number of signals in each path by simply going through the node sequence for the path. For simplicity, the data on signals were limited to tagging with the mere presence of signals. The data contained no detailed information such as number of phases, cycle lengths, or approach delays.

The most challenging task in the development of the Python scripts for this research was the development of a procedure that identifies left or right turning movements along a given path. The procedure performs a geometric calculation on nodes and the links along the paths (real or computed) under examination.

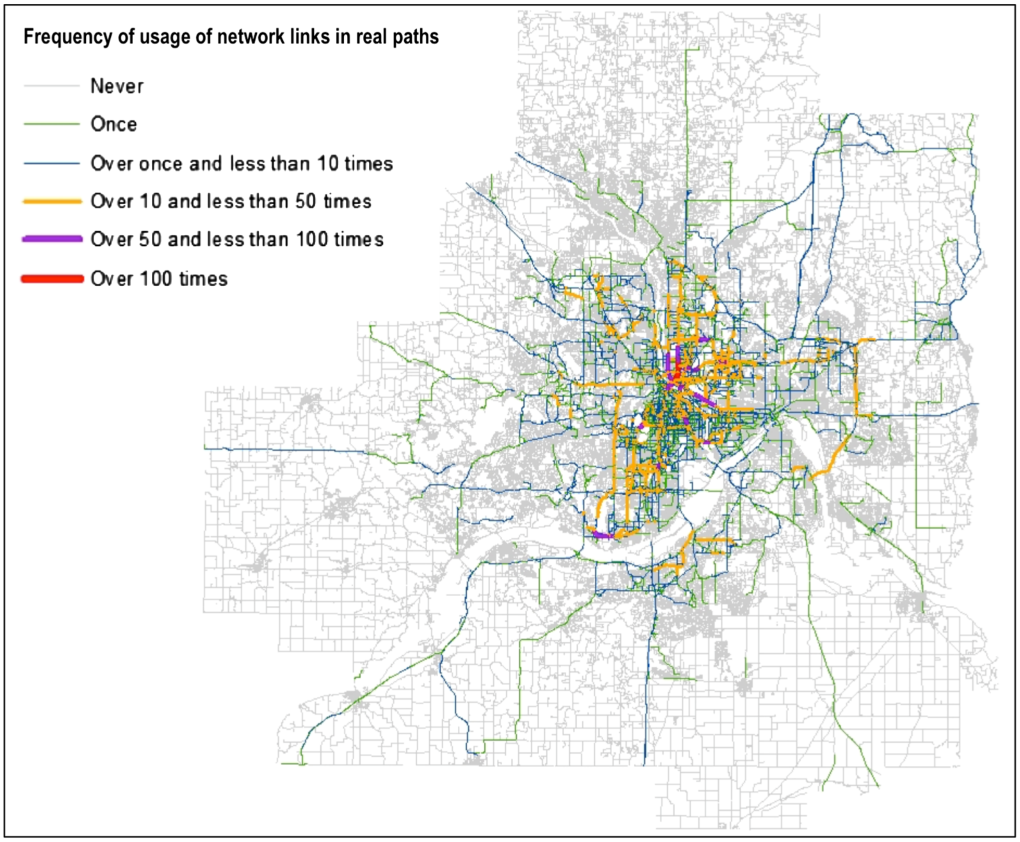

Fig. 3. The Twin Cities street network and trips trajectories used in the study

A turning movement was identified when the angle between two successive links was greater than 45°. As shown in Fig. 4, this angle θ is the absolute value of the difference by subtracting angle α from angle β. As the figure illustrates, angle α and angle β are angles between links and the positive direction of the x-axis. Using coordinates of the three nodes A, B, and C, which form the two links, angles α and β can be calculated using the equations shown. Starting from the first node in the path sequence, for each three-node group the procedure computed the angle θ. If θ was greater than 45°, a turning movement was recognized to occur at node B. Depending on the direction of the trip and the turn, the movement was flagged as a left or right turn. The procedure's logic for identifying the left or right turn is also illustrated in Fig. 4.

The procedure went through the whole path-link sequence to count the number of left and right turns along the paths studied. Because the lengths of the paths varied, the counts of signals and turns along the path were normalized as the number of signals or turns per 1.61 km (per 1 mi) for later comparisons.

Turn Penalties

One of the key determinants of path selection process is delay associated with turns along the way. The delays can be a result of the presence of signals and/or turn maneuvers. Search algorithms have to account for these turn penalties to find realistic paths between any given origin and destination. Literature on the development and implementation of these algorithms has been fairly limited. Very few studies have focused on appropriate values for turn penalties and on how to quantify these penalties. Thériault et al. (1999) chose penalties of 24 s for left turns and 12 s for right turns. In another study, 30 and 7.5 s were thought more suitable penalties for left and right turns, respectively (Yiannakoulias et al. 2013). Using the GPS-tracked path trajectory data used in this study, a methodology was developed to ascribe turn penalties for signalized and nonsignalized intersections (Zhou 2014, Chapter 7). Table 1 presents a consolidated summary of turn penalties obtained by applying this methodology to the path trajectory data. Shortest path search algorithms take into account the results of this analysis.

Road Classes along Paths

For analytical convenience, all the 23 road classifications in the network identified by census feature class codes (CFCC) were combined into three major classes: primary roads (CFCC codes A1 and A2); secondary roads (CFCC code A3); and local roads (CFCC codes A4, A51, and A73). In terms of the number of links and the total link length, the class of local roads is dominant in the network. The primary road class has values generally higher than, but comparable to, the values for secondary road class (Zhou 2014).

To exclude the possibility that drivers have to choose a certain road class merely because of the composition of the network rather than their preferences, a concept of *road availability* was defined. This concept reflects the degree to which a specific class of facility can be chosen by drivers when they are making a trip. The road availability measure cannot be applied to the entire network, because only a portion of the network is relevant for a specific trip. This relevant portion, defined as *trip proximity*, is the smallest rectangular area able to cover the trip path (Fig. 5). This rectangular area was deemed appropriate for trip proximity based on one of the triangle laws of inequality. The law states that the length of each side is less than the sum of the lengths of the other two sides and greater than the difference between these two lengths. In Fig. 5, the chosen path by the driver divides trip proximity into two approximately equal triangles.

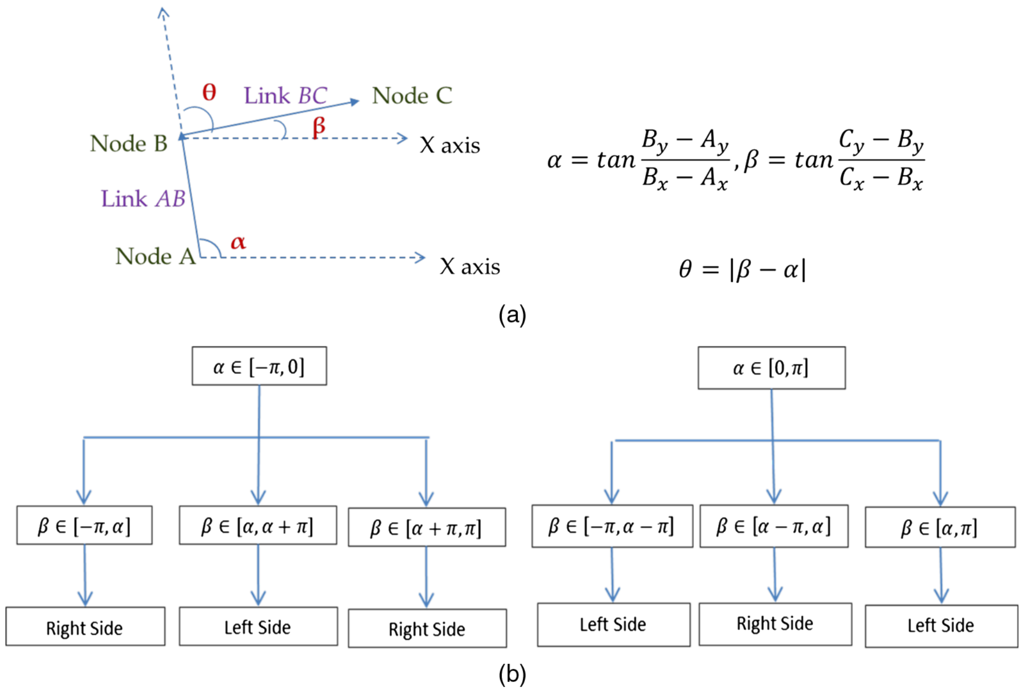

Fig. 4. Identifying turns along a path: (a) measuring turning angles; (b) identifying the direction (left or right) of a turn

The portion of a trip represented in each road class was first computed. Total lengths of all link segments along all paths were then computed for each road class within the area of trip proximity. The following two measures were computed from these accumulation counters:

1. Percentage of trip length in each road class; and
2. Percentage of trip length in each road class normalized by length of road class within the trip proximity.

The Python script examined link sequences of each path and classified links, summing up the link lengths by CFCC. The total length of each roadway class was divided by total path lengths to obtain usages for this class. The comparison between availability and usage of a specific trip can rule out the possible effect of the network composition on path selection, therefore helping to identify drivers' preferences among various road classifications.

Table 1. Average Turn Penalties from Path Trajectories

Turning movements	Number of turns	Average turn penalties (s)
Signalized left turn	3,299	29.41
Signalized right turn	2,960	15.06
Nonsignalized left turn	7,105	14.71
Nonsignalized right turn	5,465	12.51

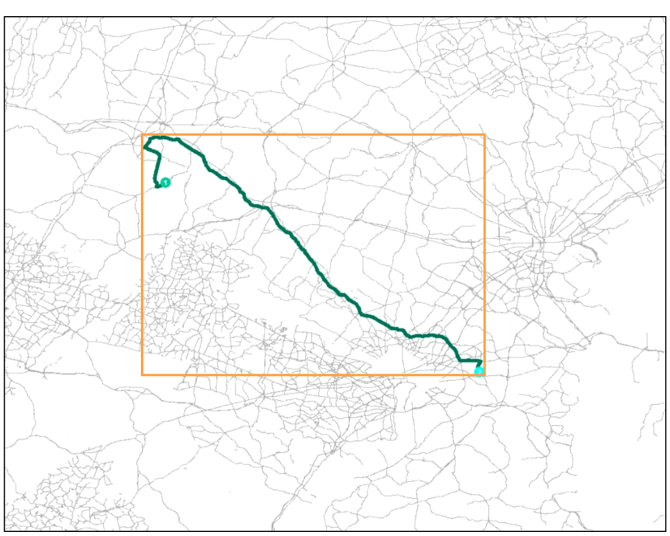

Fig. 5. Area of trip proximity

Effect of Network Characteristics on Path Choice

For analytical conveniences, real paths, SDPs, and STPs were categorized into four groups by their length. On real paths, the median value of the path length was 2.47 km (1.54 mi) and the average length of all paths was 3.51 km (2.18 mi). Thus, it was reasonable to put all paths shorter than 1.61 km (1 mi) into one group and paths between 1.61 and 8.05 km (1 and 5 mi) in another. The maximum path length in the data set was 49.17 km (30.55 mi) and relatively long paths represented only a small percentage of the trips. For this reason, all paths longer than 16.1 km (10 mi) were categorized as one group. Therefore, the data on real paths were presented in four groups:

1. Less than 1.61 km (1 mi);
2. 1.61–8.05 km (1–5 mi);
3. 8.05–16.10 km (5–10 mi); and
4. Longer than 16.1 km (10 mi).

Shortest distance and shortest time paths were categorized in the same way.

As illustrated in Fig. 6, most trips (3,721 or 65%) had a path length 1.61–8.05 km (1–5 mi). With paths shorter than 1.61 km (1 mi) together, over 90% of trips had a path length less than 8.05 km (5 mi), whether the path was real or computed. One of the reasons for relatively fewer longer trips in the data set was that

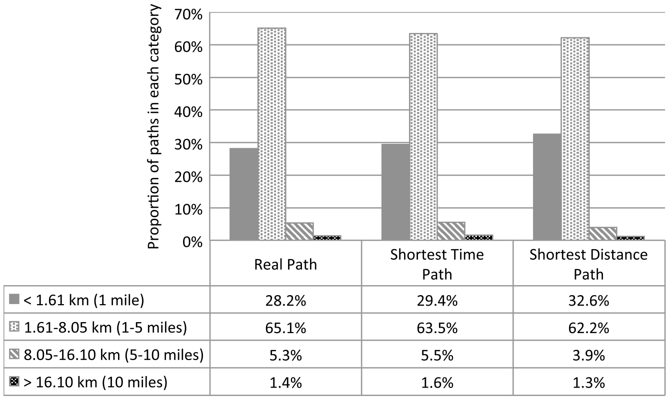

Fig. 6. Composition of path-length distributions

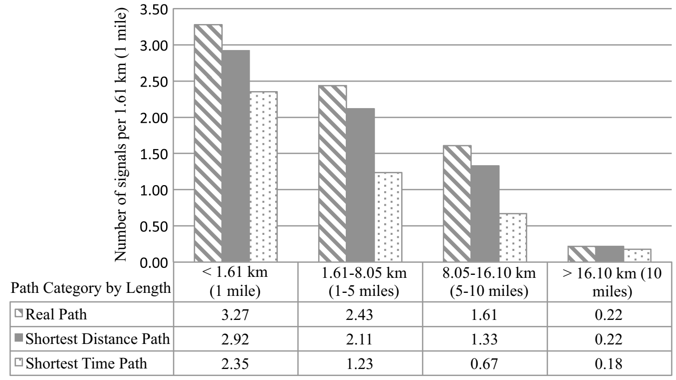

Fig. 8. Number of signals per 1.61 km (1 mi)

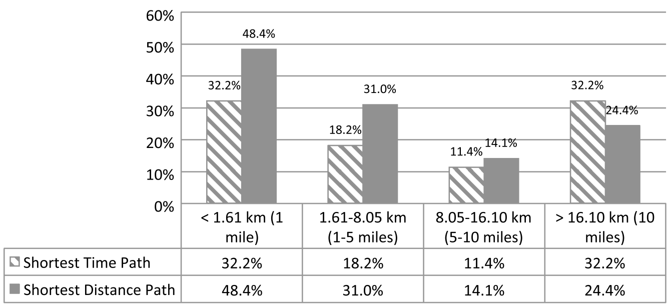

Fig. 7. Computed paths compared to real paths by path length

some of the longer trips were removed in previous steps because they were repetitive commuting trips made by the same driver.

About 35% (1,997) shortest distance paths and 22% (1,266) shortest time paths were found to be identical to their real path counterparts. Fig. 7 further illustrates the breakdown of shortest paths that were identical to corresponding real paths in each path length category. The highest identical rates occurred on paths with shorter lengths, for both shortest distance and shortest time paths. The rate decreases when the path length becomes longer. The category of 8.05–16.10 km (5–10 mi) had the lowest identical rate. The identical rate analysis shows that drivers did not necessarily choose the SDPs or the STPs. This finding means that drivers must be influenced by other factors when they choose their real paths.

As shown in Table 2, both real and computed paths have average path lengths comparable to each other. The two computed paths, SDP and STP, had the same average length for paths less than 8.05 km (5 mi). Real paths had average length slightly longer than theoretical paths (SDP and STP), except for the trips between 8.05 and 16.10 km (5 and 10 mi). In this category, STPs had the longest average path length. For path lengths more than 16.10 km (10 mi), SDPs took the longest time. For all trips, STPs took less time than the other two types, which was consistent with what was expected. It is clearly evident from the table that in making a route choice, drivers are willing to spend longer times or travel longer distances than the time and distance via the STP or SDP, respectively.

As mentioned earlier, the volunteer trip makers were local to Twin Cities and, therefore, it was assumed that the trip makers had prior knowledge about the network variables along their chosen and alternative paths. It was further assumed that for each trip the trip maker's choice set included only three paths: the actual path chosen, SDP, and STP. If the real path was a deviation from the SDP or STP, the following network variables along the paths were able to explain the deviation:

- Presence of signals;
- Overall number of turning movements;
- Number of turning movements at signalized intersections;
- Number of turning movements at nonsignalized intersections; and
- Class of roadway links.

It was also assumed that left and right turns have different effects on path choice and were therefore treated separately.

In order to identify the influence of these network variables on path choice, a controlled statistical test was necessary. A paired t-test, which was used to compare two population means via two samples in which observations in one sample could be paired with observations in the other sample, suited this need. A series of paired t-tests were performed, which paired sample means of real paths versus SDPs and real paths versus STPs.

Effect of Signals

Fig. 8 illustrates that both SDPs and STPs had fewer signals than real paths when the length was shorter than 16.1 km (10 mi). The paired t-test results shown in Table 4 are consistent with this observation. Most of the p-values in Table 3 are more than 0.05,

Table 2. Average Path Length and Path Times

		Path length category			
Average path	Path	<1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
Length in km (mi)	Real paths	0.74 (1.19)	2.20 (3.54)	6.36 (10.24)	15.01 (24.17)
	Shortest distance paths	0.69 (1.11)	2.16 (3.48)	6.41 (10.32)	15.00 (24.15)
	Shortest time paths	0.69 (1.11)	2.20 (3.54)	6.50 (10.47)	14.58 (23.47)
Time in minutes	Real paths (GPS times)	7.81	11.13	18.33	27.70
	Shortest distance paths	1.33	3.81	10.17	18.46
	Shortest time paths	2.61	5.81	11.92	17.30

Table 3. Effect of Signals on Path Choice

Path	Statistic	Path length category <1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
Real path	Mean signals per mile along real path (μ_r)	3.304	2.540	1.644	0.245
	Degrees of freedom	1,611	3,721	304	78
Real path versus shortest distance path	Mean signals per mile along SDP (μ_s)	3.059	2.247	1.471	0.252
	Mean difference (μ_r–μ_s)	0.236	0.288	0.111	−0.092
	t-statistic	7.390	16.137	2.290	−2.423
	p-value (one-tailed)	1.000	1.000	0.989	**0.009**
Real path versus shortest time path	Mean signals per mile along STP (μ_s)	2.369	1.359	0.591	0.167
	Mean difference (μ_r–μ_s)	0.935	1.181	1.053	0.078
	t-statistic	15.815	40.608	11.645	2.037
	p-value (one-tailed)	1.000	1.000	1.000	**0.023**

Note: H_0: (μ_r–μ_s) ≥ 0, H_a: (μ_r–μ_s) < 0; at 95% confidence interval of the difference, the p-value less than 0.05 indicates that H_0 may be rejected (or H_a may be accepted); values in bold indicate that real paths have statistically fewer signals per mile than do shortest paths.

which means the null hypothesis that real paths have more signals than shortest paths cannot be rejected. The only two exceptions were in the comparison to SDPs and STPs for trips longer than 16.1 km (10 mi). Considering the small portion of long paths, these p-values could be regarded as the result of outliers. Therefore, it can be concluded that the number of signals along the path was not an influential factor that significantly affected drivers' path choice. The test indicated that the opposite was true for trips shorter than 16.1 km (10 mi). That is, drivers embraced more signals along their paths than the number of signals that were present along the SDP or STP.

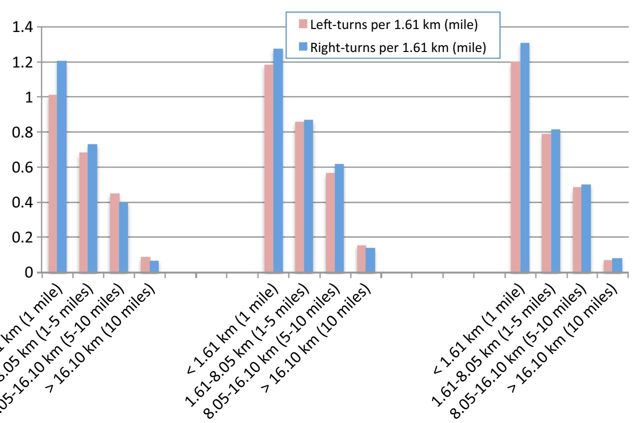

Fig. 9. Compositions of turns in real, shortest distance and shortest time paths

Effect of Turns Regardless of Signal Status

Fig. 9 illustrates the number of turns per 1.61 km (1 mi) along paths for real paths, for SDPs, and for STPs. It can be generally ascertained that the longer the path length, the fewer the number of turns. This pattern was the same for all sets and for both left and right turns.

Real paths clearly had lower average turns per 1.61 km (1 mi) in all length categories than computed paths. Furthermore, 52.7% of all turns in real paths were right turns and for trips shorter than 8.05 km (5 mi), the percentage of right turns increased to 54% (Table 5). Also, all paths had fewer left turns than right turns on average for trips shorter than 8.05 km (5 mi). These two observations are particularly important because they lead to the following questions:

1. Do real paths tend to have fewer left turns than right turns for shorter trips?
2. For shorter trips, do real paths tend to have fewer overall turns than the corresponding shortest distance path or shortest time path?

Presented in Table 4 are the results of the paired t-test analysis, which addresses the first question. In this analysis, the number of left turns per 1.61 km (1 mi) was paired against the number of right turns per 1.61 km (1 mi) for each of the nearly 6,000 real-world trips in the data set.

The analysis in Table 4 indicates that for trip lengths 8.05 km (5 mi) or shorter, routes chosen by drivers will have more right turns than left turns. This observation statistically confirms the intuition that drivers tend to minimize or avoid left turns. Furthermore, it also validates the case for a vehicle maintenance recommendation that is unrelated to route choice: tires must be rotated at a regular frequency because the majority of the turns drivers make are right turns.

The results of paired t-tests in Table 5 address the second question posed earlier. The t-tests indicated that in all path length

Table 4. Left Turns versus Right Turns in Real Paths

Statistic	Path length category <1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
Degrees of freedom	1,611	3,721	304	78
Mean left turns per 1.61 km (1 mi) along real Path (μ_{lt})	1.011	0.678	0.468	0.101
Mean right turns per 1.61 km (1 mi) along real path (μ_{rt})	1.211	0.729	0.413	0.080
Mean difference (μ_r–μ_s)	−0.199	−0.051	0.055	0.022
t-statistic	−4.839	−3.570	1.958	2.242
p-value (two-tailed)	**0.000**	**0.000**	0.974	0.986

Note: H_0: (μ_{lt}–μ_{rt}) ≥ 0, H_a: (μ_{lt}–μ_{rt}) < 0; at 95% confidence interval of the difference, the p-value less than 0.05 indicates H_0 may be rejected and H_a may be accepted; values in bold indicate that real paths have statistically fewer left turns than right turns.

Table 5. Effect of Turns on Path Choice

Path	Statistic	Path length category <1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
All left turns (signalized or nonsignalized)					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean left turns per 1.61 km (1 mi) along real path (μ_r)	1.011	0.678	0.468	0.101
Real path versus shortest distance path	Mean left turns per 1.61 km (1 mi) along SDP (μ_s)	1.182	0.859	0.566	0.155
	Mean difference (μ_r–μ_s)	−0.171	−0.181	−0.118	−0.067
	t-statistic	−5.570	−12.818	−3.485	−2.839
	p-value (one-tailed)	**0.000**	**0.000**	**0.000**	**0.003**
Real path versus shortest time path	Mean left turns per 1.61 km (1 mi) along STP (μ_s)	1.133	0.823	0.395	0.092
	Mean difference (μ_r–μ_s)	−0.121	−0.144	0.073	0.009
	t-statistic	−2.614	−8.376	1.130	1.068
	p-value (one-tailed)	**0.005**	**0.000**	0.870	0.856
All right turns (signalized or nonsignalized)					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean right turns per 1.61 km (1 mi) along real path (μ_r)	1.211	0.729	0.413	0.080
Real path versus shortest distance path	Mean right turns per 1.61 km (1 mi) along SDP (μ_s)	1.272	0.865	0.616	0.138
	Mean difference (μ_r–μ_s)	−0.066	−0.136	−0.219	−0.067
	t-statistic	−2.095	−10.802	−7.877	−2.839
	p-value (one-tailed)	**0.018**	**0.000**	**0.000**	**0.003**
Real path versus shortest time path	Mean right turns per 1.61 km (1 mi) along STP (μ_s)	1.190	0.835	0.433	0.112
	Mean difference (μ_r–μ_s)	0.021	−0.106	−0.020	−0.033
	t-statistic	0.442	−6.910	−0.330	−2.909
	p-value (one-tailed)	0.671	**0.000**	0.371	**0.002**

Note: H_0: $(\mu_r–\mu_s) \geq 0$, H_a: $(\mu_r–\mu_s) < 0$; at 95% confidence interval of the difference, the p-value less than 0.05 indicates that H_0 may be rejected (or H_a may be accepted); values in bold indicate that real paths have significantly fewer left or right turns per 1.61 km (1 mi) than do shortest paths.

Table 6. Effect of Signalized Turns on Path Choice

Path	Statistic	Path length category <1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
Signalized left turns					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean signalized left turns per 1.61 km (1 mi) along real path (μ_r)	0.383	0.250	0.154	0.018
Real path versus shortest distance path	Mean signalized left turns per 1.61 km (1 mi) along SDP (μ_s)	0.285	0.197	0.089	0.014
	Mean difference (μ_r–μ_s)	0.091	0.053	0.060	−0.002
	t-statistic	5.371	8.118	6.194	−0.404
	p-value (one-tailed)	1.000	0.999	1.000	0.344
Real path versus shortest time path	Mean signalized left turns per 1.61 km (1 mi) along STP (μ_s)	0.215	0.170	0.052	0.023
	Mean difference (μ_r–μ_s)	0.168	0.080	0.102	−0.004
	t-statistic	4.592	7.100	1.902	−0.755
	p-value (one-tailed)	1.000	1.000	0.971	0.226
Signalized right turns					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean signalized right turns per 1.61 km (1 mi) along real path (μ_r)	0.426	0.275	0.152	0.022
Real path versus shortest distance path	Mean signalized right turns per 1.61 km (1 mi) along SDP (μ_s)	0.289	0.199	0.107	0.016
	Mean difference (μ_r–μ_s)	0.137	0.074	0.037	0.003
	t-statistic	7.770	12.207	2.877	0.520
	p-value (one-tailed)	1.000	1.000	0.998	0.698
Real path versus shortest time path	Mean signalized right turns per 1.61 km (1 mi) along STP (μ_s)	0.235	0.169	0.047	0.023
	Mean difference (μ_r–μ_s)	0.192	0.105	0.105	0.000
	t-statistic	5.301	9.653	1.912	−0.082
	p-value (one-tailed)	**0.000**	1.000	0.972	0.468

Note: H_0: $(\mu_r–\mu_s) \geq 0$, H_a: $(\mu_r–\mu_s) < 0$; at 95% confidence interval of the difference, the p-value less than 0.05 indicates that H_0 may be rejected (or H_a may be accepted); values in bold indicate that real paths have significantly fewer signalized left or right turns per 1.61 km (1 mi) than do shortest paths.

categories, the null hypothesis that real paths tend to have more left turns may not be rejected. Therefore, the statistical evidence suggests that real paths tend to have fewer overall turns than the corresponding shortest distance path or shortest time path. The same cannot be said about right turns.

Effect of Turns at Signalized Intersections

The results of the analysis of turns at signalized intersections along real paths, SDPs, and STPs are shown in Table 6. As the table shows, for all path lengths, mean signalized left turns along real paths were consistently higher than mean signalized left turns along

SDPs and STPs. The same observation is true for signalized right turns on routes longer than 1.61 km (1 mi). Paired t-tests also confirm that there is statistical evidence that drivers prefer a route where turns happen at signalized intersections. This evidence further supports the notion floated in the title of this research paper: that drivers embrace signals when making a path choice.

Effect of Turns at Nonsignalized Intersections

When the normalized number of nonsignalized intersections along the real path, STP, and SDP were compared in each distance category, real paths were observed to have fewer nonsignalized right turns than the theoretical paths (Table 7). The average number of turns per 1.61 km (1 mi) at nonsignalized intersections was significantly more than that of STPs for path lengths longer than 8.05 km (5 mi).

Similar to the findings for all turns regardless of signal status, nonsignalized turns per 1.61 km (1 mi) along real paths were fewer than nonsignalized turns per 1.61 km (1 mi) along STP and SDP intersections when the path length was shorter than 8.05 km (5 mi). A reasonably simple way to summarize this analysis is that there is statistical evidence to indicate that, in choosing paths for trips shorter than 8.05 km (5 mi), drivers tend to minimize turns at nonsignalized intersections.

Effect of Road Classification

Fig. 10 compares road availability to road usage of real paths for four path lengths and three major road classes. Local roads within proximity dominate for all length groups. Although the combined availability of primary and secondary roads was only 10% or less, these two higher classes of roads accounted for a disproportionally larger portion of the road usage. Regardless of the network composition, drivers were willing (not forced) to choose roads with a higher level of functional class.

As would be expected, as path lengths increased, the portion of the primary road in a real path became larger, and the portion of the local road became smaller. The percentage of secondary roads in real paths first increased with path length and then decreased. For paths shorter than 16.1 km (10 mi), over 50% usage belonged to local roads, and primary roads only took the smallest portion. Only when trip length was longer than 16.1 km (10 mi) did the percent share of local roads become lower than the other two classes. The percent share of primary roads was the highest for trips longer than 16.1 km (10 mi).

Primary and Secondary Road Usage

This section further explores commonalities and differences between real paths and shortest paths in terms of usage of primary and secondary roads. In each road class, SDPs and STPs trended the same way as real paths when path lengths increased (Fig. 11). In other words, usage of primary roads increased while that of local roads decreased. The highest usage of secondary roads happened in paths between 8.05 and 16.1 km (5 and 10 mi). However, for trips longer than 16.1 km (10 mi), the smallest percentage was secondary roads for the SDP. For trips between 8.05 and 16.1 km (5 and 10 mi), the largest percentage was secondary roads for the STP.

Primary, secondary, and local road usage statistics on real paths were compared to those of STPs and SDPs. On average, STPs occurred more along primary and secondary roads than did the other two sets, and by extension, fewer STPs occurred along local roads for all trips. This is reasonable because primary and secondary roads offer higher average traveling speeds and less travel time accordingly. Furthermore, for the shortest time path, the percentage of local roads declined more sharply than for the other two types of

Table 7. Effect of Nonsignalized Turns on Path Choice

Path	Statistic	Path length category <1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
Nonsignalized left turns					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean nonsignalized left turns per 1.61 km (1 mi) along real path (μ_r)	0.628	0.428	0.313	0.083
Real path versus shortest distance path	Mean nonsignalized left turns per 1.61 km (1 mi) along SDP (μ_s)	0.897	0.663	0.477	0.141
	Mean difference (μ_r–μ_s)	−0.262	−0.234	−0.178	−0.065
	t-statistic	−8.757	−17.707	−5.669	−2.960
	p-value (one-tailed)	**0.000**	**0.000**	**0.000**	**0.002**
Real path versus shortest time path	Mean nonsignalized left turns per 1.61 km (1 mi) along STP (μ_s)	0.918	0.652	0.342	0.080
	Mean difference (μ_r–μ_s)	−0.289	−0.224	−0.029	0.003
	t-statistic	−9.323	−17.049	−1.122	0.329
	p-value (one-tailed)	**0.000**	**0.000**	0.131	0.628
Nonsignalized right turns					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean nonsignalized right turns per 1.61 km (1 mi) along real path (μ_r)	0.784	0.454	0.261	0.058
Real path versus shortest distance Path	Mean nonsignalized right turns per 1.61 km (1 mi) along SDP (μ_s)	0.983	0.667	0.509	0.122
	Mean difference (μ_r–μ_s)	−0.203	−0.210	−0.256	−0.075
	t-statistic	−6.681	−16.465	−9.581	−3.396
	p-value (one-tailed)	**0.000**	**0.000**	**0.000**	**0.000**
Real path versus shortest time path	Mean nonsignalized right turns per 1.61 km (1 mi) along STP (μ_s)	0.955	0.665	0.386	0.090
	Mean difference (μ_r–μ_s)	−0.171	−0.211	−0.125	−0.032
	t-statistic	−5.109	−17.608	−6.059	−3.105
	p-value (one-tailed)	**0.000**	**0.000**	**0.000**	**0.001**

Note: H_0: $(\mu_r–\mu_s) \geq 0$, H_a: $(\mu_r–\mu_s) < 0$; at 95% confidence interval of the difference, the p-value less than 0.05 indicates that H_0 may be rejected (or H_a may be accepted); values in bold indicate that real paths have significantly fewer nonsignalized left or right turns per 1.61 km (1 mi) than do theoretical shortest paths.

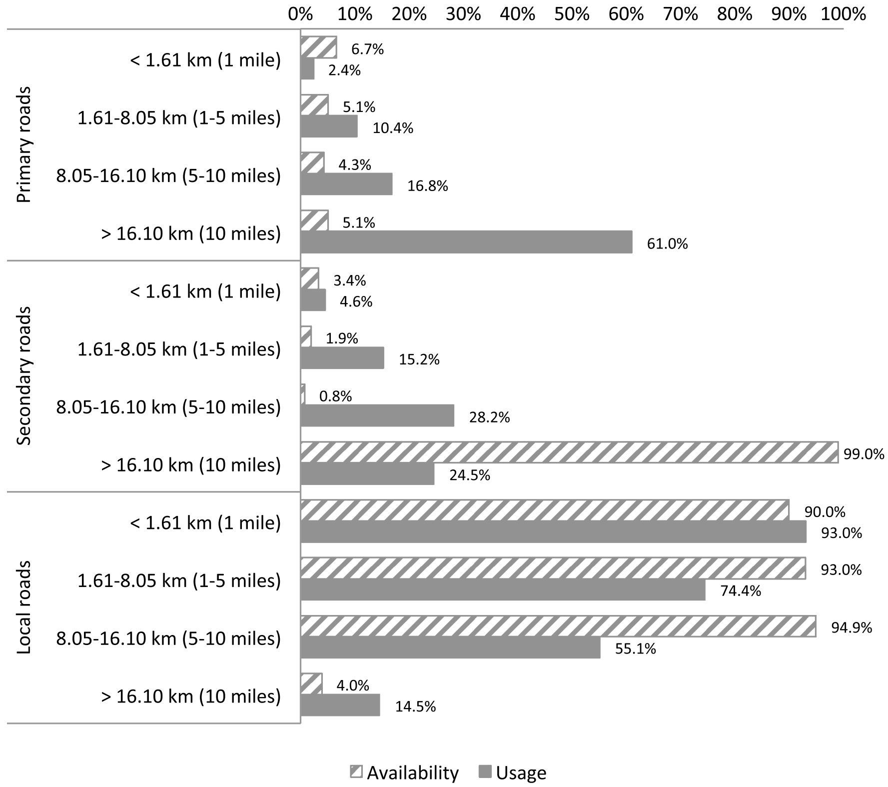

Fig. 10. Availability versus usage of various road classes in real paths

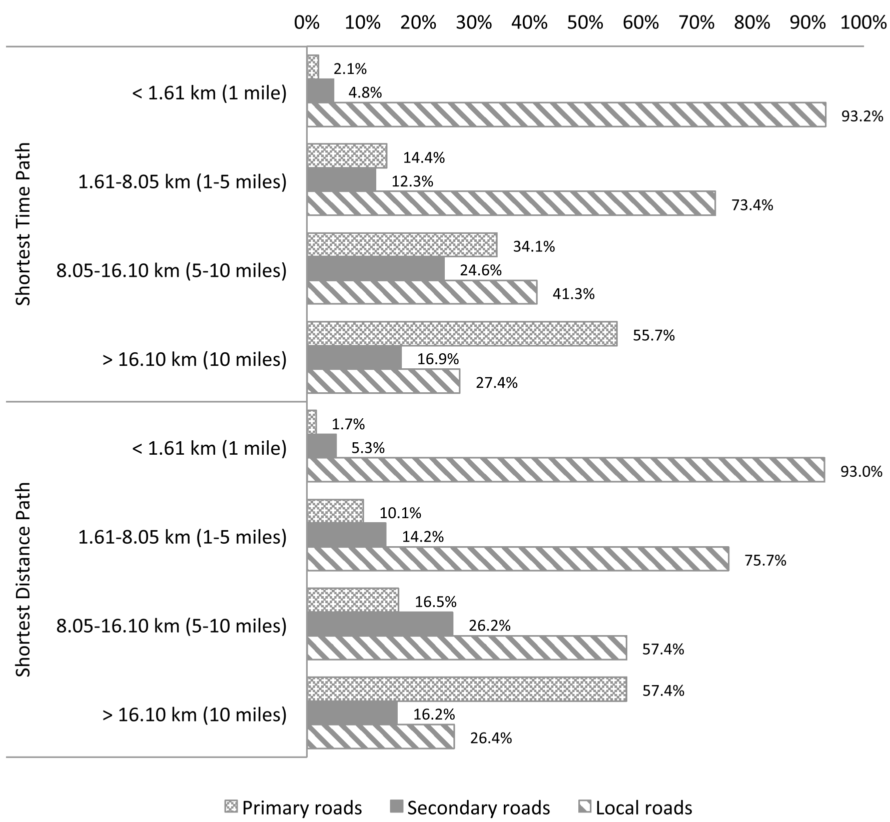

Fig. 11. Road class distribution in SDP and STP

Table 8. Effect of Primary Road on Path Choice

Path	Statistic	Path length category			
		<1.61 km (1 mi)	1.61–8.05 km (1–5 mi)	8.05–16.10 km (5–10 mi)	>16.10 km (10 mi)
Usage of primary roads in paths					
Real path	Degrees of freedom	1,611	3,721	304	78
	Mean percentage of primary road along real path (μ_r)	0.011	0.078	0.147	0.512
Real path versus shortest distance path	Mean percentage of primary roads along SDP (μ_s)	0.013	0.088	0.128	0.540
	Mean difference (μ_r–μ_s)	0.0080	0.005	0.037	0.062
	t-statistic	3.667	3.871	3.726	3.534
	p-value (one-tailed)	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Real path versus shortest time path	Mean percentage of primary roads along STP (μ_s)	0.023	0.119	0.301	0.536
	Mean difference (μ_r–μ_s)	−0.012	−0.041	−0.154	−0.024
	t-statistic	−5.503	−16.176	−10.480	−1.585
	p-value (one-tailed)	**0.000**	**0.000**	**0.000**	0.059
Usage of secondary roads in paths					
Real path	Degrees of freedom	1,610	3,721	304	78
	Mean percentage of secondary road along real path (μ_r)	0.030	0.121	0.243	0.220
Real path versus shortest distance path	Mean percentage of secondary road along SDP (μ_s)	0.038	0.131	0.230	0.187
	Mean difference (μ_r–μ_s)	0.002	0.008	0.044	0.074
	t-statistic	1.162	3.352	3.565	3.985
	p-value (one-tailed)	0.123	**<0.001**	**<0.001**	**<0.001**
Real path versus shortest time path	Mean percentage of secondary road along STP (μ_s)	0.042	0.116	0.246	0.200
	Mean difference (μ_r–μ_s)	−0.012	0.005	−0.003	0.020
	t-statistic	−4.180	1.593	−0.215	1.438
	p-value (one-tailed)	**0.000**	0.944	0.415	0.923

Note: H_0: $\mu_r–\mu_s \leq 0$, H_a: $\mu_r–\mu_s > 0$; at 95% confidence interval of the difference, the p-value less than 0.05 indicates it is safe to reject H_0 and H_a may be accepted; values in bold indicate that real paths have significantly higher percentage of primary or secondary roads than the shortest paths.

path, with increasing path length. Meanwhile, SDPs used fewer primary roads than real paths in all categories.

Paired sample t-tests showed that real paths had higher percentages of primary and secondary roads than the shortest distance paths in almost all circumstances (Table 8). Only for paths shorter than 1 mi did real paths have less usage of secondary roads than the shortest distance paths, where the p-value (0.123) was more than the level of significance (0.05). Proportion of primary roads in real paths was higher than the proportion in SDPs or STPs for the same O-D pairs.

Paired t-tests also indicated that real paths gravitated more toward the higher usage of both primary and secondary roads than theoretical paths computed using the shortest time. On the other hand, there was no significant difference between real paths and SDPs with regards to usage of roads of higher classes.

However, real paths did not have more primary and secondary roads than shortest time paths. This confirms the intuition that people prefer major roads because they have higher posted speed limits and fewer interruptions, which leads to less travel time. For this reason only there was no significant difference between real paths and shortest time paths regarding the distribution of road classifications.

For shorter trips, the majority of road usage was in the local road category. Local roads have more intersections and consequently may generate more turning movements. However, there was no obvious pattern for trips of more than 8.05 km (5 mi) because on these trips the influence of turns declined because of the increasing percentage of primary and secondary roads.

Conclusions

This research examined the influence of street network and turn variables on drivers' path choices in a major metropolitan area, Minneapolis-St. Paul, Minnesota. Controlling for these variables, real-world path trajectories were compared to the computed shortest distance and shortest time paths. The analysis indicated that drivers are willing to spend longer time or travel longer distances on paths that have fewer turning movements. The study confirms the intuition that drivers tend to avoid turning movements during their travel. There is statistical evidence to indicate that real paths have fewer turns per 1.61 km (1 mi) than both shortest time and shortest distance paths. When they must make a left turn or right turn to complete their trips, drivers seem more prone to making the turn at a signal-controlled intersection, while at the same time trying to minimize the number of turns occurring at nonsignalized intersections. Most notably, for trips shorter than 8.05 km (5 mi) in length, real paths have a statistically significant fewer left turns per 1.61 km (1 mi) than right turns per 1.61 km (1 mi). This leads to the conclusion that drivers tend to minimize left turns while selecting a path.

The research also presented statistical evidence that the number of traffic signals along the chosen route versus its theoretical counterparts is not a significant factor in path-choice processing. Statistical analyses also revealed that, in terms of the number of signals per mile, real paths contain more signals than their theoretical paths. Prior to this effort, most studies relating path-choice behavior to network and path characteristics were based on stated preference surveys. This research used a large data set of paths with trajectories tracked by GPS to identify the effects of certain network and path characteristics on drivers' route choice. Compared to stated preference surveys, the GPS tracking data are a better representation of how drivers choose routes in actual practice.

Study findings confirm, with statistical evidence, some intuitive and some not-so-intuitive expectations and hypotheses. The results of this study may be generalized with caution for route-choice behavior of drivers who are familiar with the characteristics of the street network, signals, and traffic patterns. The methodologies used in this research will make it easier to find paths more consistent with drivers' real choices and consequently provide more sound and solid solutions to traffic assignment problems and other problems in transportation planning. Therefore, the main contribution from this study may be seen as the methodologies that were developed to address the research questions, rather than the end results. By applying these methods to a wide range of vehicle trajectory data sets currently available, thanks to the prevalence of smart phones and advanced vehicle technologies, the effect of network variables on path choice behavior can be clearly understood. Similar studies done for different urban areas will lead to developing solutions to network-management and transportation-planning problems at a local jurisdictional level. A collective assessment of similar studies using different data sets will lead to developing philosophical underpinnings of groundbreaking theories similar to the Wardrop criteria, which led to the development of user equilibrium assignment.

While the trends observed from the data used in this research may be applicable to other population groups and geographical areas comparable to the study area, care must be taken in generalizing the results or using the numerical values in traffic-effect and transportation-planning studies. Certain observations made in this study may be applicable only to the trip data collected from 44 volunteers in Minneapolis-St. Paul, Minnesota. More real-world observation data from other metropolitan areas and rural areas would provide more statistically significant results and findings.

It is not clear if any of the path choices in the data set were influenced by real-time travel information received prior to or during the trip. However, the control variables of this study were network and turn characteristics, which are independent of travel information. Therefore, the findings of this study are independent of the actions taken by the drivers based on any real-time travel information. As with any study, expansion of the data available for analysis would improve the significance and robustness of the results of this research.

Acknowledgments

This study was partially supported by a grant from the U.S. Department of Transportation under the University Transportation Centers program. The authors thank Dr. Mark Houck and Dr. Chaowei Yang of George Mason University for their help in reviewing the research work and offering valuable suggestions during the course of this study.

References

Abdel-Aty, M. A., Kitamura, R., and Jovanis, P. P. (1995). "Investigating effect of travel time variability on route choice using repeated-measurement stated preference data." *Trans. Res. Rec.*, 1493, 39–45.

Antonisse, R., Daly, A., and Ben-Akiva, M. (1989). "A highway assignment method based on behavioral models of car drivers' route choice." *Transp. Res. Rec.*, 1220, 1–11.

ArcGIS 10.3 [Computer software]. ESRI, Redlands, CA.

Ben-Akiva, M., and Bierlaire, M. (1999). "Discrete choice methods and their applications to short term travel decisions." *Handbook of transportation science*, Springer, 5–33.

Ben-Akiva, M., and Bolduc, D. (1996). "Multinomial probit with a logit kernel and a general parametric specification of the covariance structure." Departement d'economique, Universite laval with Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Bernstein, D., and Kornhauser, A. (1996). "An introduction to map-matching for personal navigation assistants." Princeton Univ., Princeton, NJ.

Cascetta, E., Nuzzolo, A., Russo, F., and Vitetta, A. (1996). "A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks." *Proc., 13th Int. Symp. on Transportation and Traffic Theory*, Pergamon, Lyon, France, 697–711.

Cascetta, E., and Papola, A. (1998). "An implicit availability/perception random utility model for path choice." *Proc., TRISTAN III*, Vol. 2.

Chatterjee, A., and Venigalla, M. M. (2003). "Travel demand forecasting for urban transportation planning." *Handbook of transportation engineering*, M. Kutz, ed., McGraw-Hill, New York.

Chen, W. H., and Jovanis, P. P. (2003). "Driver en route guidance compliance and driver learning with advanced traveler information systems: Analysis with travel simulation experiment." *Transp. Res. Rec.*, 1843, 81–88.

Chu, C. (1989). "A paired combinatorial logit model for travel analysis." *Proc., 5th World Conf. on Transportation Research*, Ventura, CA, 295–309.

Correa, J. R., and Stier-Moses, N. E. (2011). "Wardrop equilibria." *Wiley encyclopedia of operations research and management science*, Wiley, New York.

Daganzo, C. F., and Sheffi, Y. (1977). "On stochastic models of traffic assignment." *Trans. Sci.*, 11(3), 253–274.

Dijsktra, E. W. (1959). "A note on two problems in connection with graphs." *Numrische Mathematik*, 1(1), 269–271.

Du, J., and Aultman-Hall, L. (2007). "Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues." *Trans. Res. Part A: Policy Pract.*, 41(3), 220–232.

Jackson, W. B., and Jucker, J. V. (1982). "An empirical study of travel time variability and travel choice behavior." *Transp. Sci.*, 16(4), 460–475.

Jan, O., Horowitz, A., and Peng, Z. (2000). "Using global positioning system data to understand variations in path choice." *Transp. Res. Rec.*, 1725, 37–44.

Li, H. (2004). "Investigating morning commute route choice behavior using global positioning systems and multi-day travel data." Ph.D. dissertation, Georgia Institute of Technology, Atlanta.

Liu, B. (1996). "Intelligent route finding: Combining knowledge, cases and an efficient search algorithm." *Proc., 12th Int. Conf. Artificial Intelligence, Cases and an Efficient Search Algorithm*, PITMAN, 380–384.

McFadden, D., and Train, K. (2000). "Mixed MNL models for discrete response." *J. Appl. Econ.*, 447–470.

Polydoropoulou, A., Ben-Akiva, M., and Kaysi, I. (1994). "Influence of traffic information on drivers' route choice behavior." *Transp. Res. Rec.*, 1453, 56–65.

Quddus, M. A., Noland, R. B., and Ochieng, W. Y. (2006). "A high accuracy fuzzy logic based map-matching algorithm for road transport." *J. Intell. Transp. Syst.*, 10(3), 103–115.

Quddus, M. A., Ochieng, W. Y., and Noland, R. B. (2007). "Current map-matching algorithms for transport applications: State-of-the-art and future research directions." *Transp. Res. Part C: Emerging Technol.*, 15(5), 312–328.

Srinivasan, K. K., and Mahmassani, H. S. (2000). "Modeling inertia and compliance mechanisms in route choice behavior under real-time information." *Transp. Res. Rec.*, 1725, 1725, 45–53.

Stern, E., Holm, E., and Maarseveen, M. V. (1993). "Information and commuters behaviour: A comparative analysis." *Eur. Move*, 141–155.

Thériault, M., Vandersmissen, M. H., Lee-Gosselin, M., and Leroux, D. (1999). "Modelling commuter trip length and duration within GIS: Application to an O-D survey." *J. Geogr. Inform. Decis. Anal.*, 3(1), 41–55.

Venigalla, M. M., Chatterjee, A., and Bronzini, M. S. (1999). "A specialized equilibrium assignment algorithm for air quality modeling." *Transp. Res. D*, 4(1), 29–44.

Vovsha, P. (1997). "The cross-nested logit model: Application to ode choice in the Tel-Aviv metropolitan area." *Trans. Res. Rec.*, 1607, 6–15.

Wardrop, J. G., and Whitehead, J. I. (1952). "Some theoretical aspects of road traffic research." *ICE Proc.: Eng. Div.*, 1(5), 767–768.

White, C. E., Bernstein, D., and Kornhauser, A. L (2000). "Some map matching algorithms for personal navigation assistants." *Transp. Res. Part C: Emerging Technol.*, 8(1–6), 91–108.

Yiannakoulias, N., Bland, W., and Svenson, L. W. (2013). "Estimating the effect of turn penalties and traffic congestion on measuring spatial accessibility to primary health care." *Appl. Geogr.*, 39, 172–182.

Zhou, X. (2014). "Revealed path choice behavior and network pruning for efficient path finding." Ph.D. dissertation, George Mason Univ., Fairfax, VA.

Zhou, X., and Venigalla, M. M. (2014). "Influence of turns at signalized intersections." *Transportation Research Board Annual Meeting*, Transportation Research Board, Washington, DC.

Zhuang, L., Gong, J., He, Z., and Xu, F. (2012). "Framework of experienced route planning based on taxis' GPS data." *15th Int. IEEE Conf. on Intelligent Transportation Systems*, IEEE, New York.