Beautiful...but at What Cost? An Examination of Externalities in Geographic Vehicle Routing

I. JOHNSON, Northwestern University
J. HENDERSON, University of Minnesota
C. PERRY, University of Minnesota
J. SCHÖNING, University of Bremen
B. HECHT, Northwestern University

Millions of people use platforms such as Google Maps to search for routes to their desired destinations. Recently, researchers and mapping platforms have shown growing interest in optimizing routes for criteria other than travel time, e.g. simplicity, safety, and beauty. However, despite the ubiquity of algorithmic routing and its potential to define how millions of people move around the world, very little is known about the *externalities* that arise when adopting these new optimization criteria, e.g. potential redistribution of traffic to certain neighborhoods and increased route complexity (with its associated risks). In this paper, we undertake the first controlled examination of these externalities, doing so across multiple mapping platforms, alternative optimizations, and cities. We find, for example, that scenic routing (i.e. "beauty"-optimized routing) would remove vehicles from highways, greatly increase traffic around parks, and, in certain cases, do the same for high-income areas. Our results also highlight that the interaction between routing criteria and urban structure is complex and effects vary from city to city, an important consideration for the growing literature on alternative routing strategies. Finally, to address the lack of open implementations of alternative routing algorithms and controlled routing evaluation frameworks, we are releasing our alternative routing and evaluation platform with this paper.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**;

General Terms: Geography; Algorithms; Vehicle Routing

Additional Key Words and Phrases: externalities; alternative routing; urban structure

ACM Reference format:

Isaac Johnson, Jessica Henderson, Caitlin Perry, Johannes Schöning, and Brent Hecht. 2017. Beautiful...but at What Cost? An Examination of the Externalities in Geographic Vehicle Routing. *PACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 2, Article 15 (June 2017), 21 pages.
DOI: 10.1145/3090080

1 INTRODUCTION

The simple act of driving from one place to another is an incredibly common part of many people's lives. However, it is also a surprisingly complex task: in many cases, there are seemingly countless routes between a given origin and destination pair. While the predominant focus of the literature and applications in the geographic routing domain has historically been on minimizing travel time or distance (e.g. [3,13]), researchers and practitioners have recently shown interest in alternative routing criteria. For instance, researchers have developed routing systems that generate "scenic" routes (e.g. [41,44,59]), simpler routes (e.g. [7,47]), and safer routes (e.g. [11,21,45]), among other alternative route optimizations (e.g. [23,25,48,60]).

This work is supported in part by the National Science Foundation (CAREER IIS-1707296 and IIS-1707319) and a Lichtenberg Professorship from the Volkswagen Foundation.
Authors' addresses: I. Johnson, Northwestern University, 2240 Campus Drive, Frances Searle Building 2-431, Evanston, IL, USA, 60208; B. Hecht, Northwestern University, 2240 Campus Drive, Frances Searle Building 1-144, Evanston, IL, USA, 60208.

2474-9567/2017/6-ART15
DOI:http://dx.doi.org/10.1145/3090080

Similarly, the routing platform Waze has begun to suggest routes, at least in Rio de Janeiro, that avoid areas deemed to have higher rates of violence [39], and Microsoft owns a patent for pedestrian routing that avoids "unsafe" areas based on crime and weather factors [51].

Despite this growing interest in alternative routing strategies, there has been no controlled evaluation of the *externalities* that arise when more traditional optimization criteria such as travel time are superseded by new optimization criteria such as safety or beauty. Evaluations of these new criteria have solely considered the direct trade-off with travel time or distance, e.g. increased travel time to achieve more scenic or safer routes. These evaluations, however, miss the externalities that may arise with these new criteria, externalities that may have significant social, economic, and safety implications. For instance, at the community level, these routing approaches may lead to increased or decreased traffic in certain areas. Additionally, at the route level, these approaches may lead to routes with more turns (directly contradicting user preferences [14,27], increasing driver stress [56] and cognitive load [16,36], and potentially decreasing safety [33,39]). Moreover, anecdotal evidence suggests that these externalities may be substantial. Consider, for instance, the widespread outcry about increased traffic and noise in previously out-of-the-way neighborhoods attributed to routing algorithm changes by Waze (e.g. [18,29,57]).

In this paper, we aim to address this gap in the literature. To do so, we developed a controlled experimental routing platform and used this platform to investigate externalities that arise with three common approaches to alternative routing: scenic routing, safety routing, and simplicity routing. Specifically, examining routes from four cities – San Francisco, New York City, London, and Manilla – we ask the following research questions:

RQ1: Does optimizing on alternative criteria in routing algorithms lead to *route-level* externalities such as more complex routes?

RQ2: Does optimizing on alternative criteria in routing algorithms lead to *community-level* externalities such as increased or decreased traffic in certain areas?

Additionally, as noted above, at least one routing platform (Waze) has already implemented alternative routing techniques [28], and Microsoft has a patent on similar approaches [51]. As such, we also saw an opportunity to use our experimental platform to better understand (and track) the criteria on which popular routing platforms are optimizing and to assess whether there are any externalities associated with these criteria. Thus, in the tradition of the algorithmic auditing literature (e.g. [2,4,53]), we also asked a third research question:

RQ3: Is there evidence of alternative routing criteria being used by *popular routing platforms*? If so, what are the associated externalities?

Overall, we find that there are large externalities associated with alternative optimization criteria and that these externalities could have substantial impacts on our communities and on the nature of the routes we use. For instance, our evidence suggests that scenic routing removes vehicles from highways (where city planners generally hope to route traffic) and redirects them to parks, popular areas, and, in some cases, wealthier areas. Scenic routing also creates substantially more complex routes involving more turns and intersections, both of which are known to make routes less desirable [14,24] and are associated with negative outcomes (e.g. traffic accidents [33], decreased usability [36,56]). Additionally, safety routing creates highly local but large impacts, redistributing traffic from pre-defined banned areas to highways and surrounding thoroughfares.

Importantly, these externalities can arise even when increases in travel time appear minimal, providing evidence that externalities may be transparent under the standard paradigm for evaluating alternative routing approaches. Along the same lines, we also identified evidence that there is substantial variation in the externalities of a given algorithm across different cities and areas within a city. As such, our findings suggest that alternative routing research must involve carefully controlled evaluations across broadly diverse geographies to fully understand the costs and benefits of an algorithmic change.

The algorithmic auditing component of our work reveals that Google Maps and MapQuest likely incorporate some non-fastest-path optimizations in their routing algorithm (e.g. they generate simpler routes that spend more time on highways). However, we found no evidence (yet) of any major externalities relative to fastest-path routing, indicating, for instance, that Google Maps and MapQuest have not implemented features like Waze has in Rio de Janeiro and applied them at scale. Our methods will allow us to easily monitor this result over time to assess if this changes (e.g. if one of these routing providers begins to route people around neighborhoods with higher crime rates).

Finally, in the spirit of work such as Jurgens et al. [20], which also sought to standardize the evaluation of an algorithm family (geolocation inference algorithms), we are releasing our alternative routing and evaluation platform to facilitate improved evaluations and comparisons in this domain. In addition to allowing researchers to easily consider externalities when evaluating new routing algorithms, our platform also addresses issues in the alternative routing literature such as a lack of standards, very limited open alternative routing implementations, and inconsistent evaluation criteria. We have designed our platform to be completely open-source and easily extensible (e.g., to other optimization criteria, geographic scales, or externalities), with the hope of supporting future research and discussion of what is important in evaluating geographic vehicle routing.

2 RELATED WORK

In this section, we discuss research that motivated this work. This research emerges primarily from four areas: the large literature on alternative routing criteria, investigations into route preference, evaluation approaches in alternative routing, and the algorithms underlying commercial mapping platforms. Notably, though the research in this paper focuses on vehicle routing, we include in our discussion of related work approaches that have considered other modes of transportation as well.

2.1 Routing Using Alternative Criteria

While there is a large and growing literature on developing alternatives to shortest and fastest path routing, there has not yet been an effort to summarize this literature. As such, we conducted a survey of the literature and found that the alternative routing approaches largely fall into four categories: *positive*, *negative*, *topological*, and *personalized* (see Table 1 below for examples of each).

Table 1. A selection of categorized alternative routing papers.

CATEGORY	SUB-CATEGORIES AND PAPERS
Positive	**Scenic** (El Ali et al. 2013, McGookin et al. 2015; Quercia, Schifanella, and Aiello 2014; Runge et al. 2016; Traunmueller et al. 2013; Zhang, Kawasaki, and Kawai 2008; Zheng et al. 2013)
Negative	**High Crime** (Elsmore et al. 2014; Fu, Lu, and Lu 2014; Kim, Cha, and Sandholm 2014; Shah et al. 2011), **Weather** (Y. Li et al. 2014), **People** (Posti et al. 2014)
Topological	**Simplicity** (Duckham and Kulik 2003; Haque, Kulik, and Klippel 2006; Shao et al. 2014), **Health** (Sharker, Karimi, and Zgibor 2012), **Efficiency** (Ganti et al. 2010)
Personalized	Letchner, Krumm, and Horvitz 2006; Delling et al. 2015; Pang et al. 1995; Ziebart et al. 2008

The first two categories, positive and negative, are defined by the work of Golledge [14], which examines in part the impact of environmental features on route preferences (e.g. parks as positive, waste dumps as negative). The third category, topological, encompasses criteria such as simplicity or driving efficiency that can be derived from basic information about the road network. The final category is personalized routing, which involves learning the personal preferences of a driver (e.g. road or turn types) and designing routes that adhere to these preferences.

Within the routing algorithms literature, positive routing often takes the form of "scenic" routing. Scenic routing has been implemented in a number of ways, e.g. reweighting edges based on an assessment of the "scenicness" of their surrounding area [55], adding waypoints from scenic areas near the shortest-path route [8,44,58], by generating many paths and then choosing the most scenic [41]. Additionally, there are also examples of optimizing for other positive criteria such as "happiness" and "quiet" [41] and projects peripheral to alternative routing that propose means of sensing criteria such as desirable smells [42] for future use in routing.

Negative routing seeks to provide routes that avoid undesirable areas. Though Golledge used waste dumps as a proxy for this type of routing, the literature largely focuses on avoiding unsafe areas as defined by high incidences of violent crime [9,11,21,45]. Additional applications include avoiding dangerous weather [25] or other people [40]. Microsoft owns a patent [51] for pedestrian routing that avoids unsafe areas. Waze has already included the option to avoid high-crime areas, specifically in Rio de Janeiro, Brazil, ahead of the 2016 Olympics [39]. Waze also defaults to routing individuals around certain settlements that are off-limits to Israelis [50]. These algorithms start with the shortest path and then either add waypoints as needed to reroute away from any areas deemed undesirable or reweight edges in these areas so that they are perceived as very high cost by the algorithm.

Topological routing describes approaches that seek to optimize on some aspect of the road network itself (rather than the environment around the road network). The most common approach in the literature – other than the more traditional travel time and distance criteria – is some form of "simplicity" routing. At their core, simplicity routing approaches seek to model the ease with which a driver can follow a route, but they operationalize simplicity a number of different ways. Golledge [14], Manley et al. [27], and Li and Wu [24] modeled simplicity as minimizing the total number of turns. Algorithmic implementations of simplicity routing have taken the approach of modeling simplicity not just using turns, but as a function of the degree of each intersection and what action is taken at that intersection (i.e. go straight or turn) [7,16,47].

Finally, personalized routing approaches generally seek to learn and model an individual's route preferences, i.e. when a user usually deviates from the fastest route and, in some cases, why s/he does so [5]. These models require extensive positioning (i.e. "GPS") data from drivers in order to determine these preferences. A common approach is to learn implicit preferences for specific roads [23,60], though a recent approach by Delling et al. [5] explicitly learns the weights that each individual driver appears to give to various topological criteria (e.g., number of lanes, type of road).

We included in our experiments the most common form of alternative routing approaches in each category, with the exception of personalized routing. More specifically: for positive routing, we implemented *scenic routing*; for negative routing, we implemented *safety routing*; and for topological routing, we implemented *simplicity routing*. We did not consider personalized routing because the open "GPS" trace datasets that would be required to include personalized routing in our experiments do not exist.

2.2 Preferences for Alternative Criteria

Researchers have long known that fastest path and shortest path are not the only criteria on which people want to optimize their routes. Much of this knowledge emerges from a variety of surveys and field studies. For instance, in an influential paper in the field of geography, Golledge [14] sought to quantify the importance of various criteria in route selection. Golledge identified that minimizing distance and travel time were the most important factors, but minimizing the number of turns and maximizing the scenic/aesthetic value were also key criteria that seemed to affect which route a participant selected. Li and Wu [24] surveyed commuters in Florida and provided support for the findings of Golledge, but also determined that safety is an important criterion. Similarly, Manley et al. [27] explored which criteria best explain actual routes taken by taxi drivers in London and found that a combination of shortest distance and fewest turns was most predictive of route choice. In addition to the preference for fewer turns, increased route complexity also raises safety concerns [33] and has been directly tied to greater cognitive loads [36] and stress [56] for the driver. This literature, though sparse, further motivates our choice of the three alternative criteria that we consider in this work (beauty, safety, and simplicity).

2.3 Evaluation in Alternative Routing

As is often the case in new computing research areas (e.g. geolocation inference [20]), evaluation in the alternative routing literature is a highly heterogeneous process that makes comparisons between approaches difficult. Evaluations typically involve examining routing in 1-2 cities using a small number of routes, with the only evaluation metric being travel time or distance. Our work is the first to our knowledge that explicitly considers additional evaluation criteria. In other words, this work sheds new light on the *externalities*, or side-effects, that arise with the use of alternative routing optimization criteria. Hints to the existence and importance of these externalities come from popular media, as discussed below. Like was the case in Jurgens et al. [20] for geolocation inference, our goal in this paper is to conduct experiments that afford a more direct and nuanced comparison between approaches, enabling a more robust understanding of the externalities associated with each approach. We also examine routing in four cities with diverse geographic contexts, affording a broader view of how geography and algorithms interact that provides important new insight.

2.4 Commercial Mapping Platforms

Since MapQuest began providing online directions in 1996, most online mapping platforms have defaulted to providing the "fastest" route between a given origin and destination. The exact details of the routing algorithms being used are proprietary, often including whether or not these algorithms are optimizing on criteria other than just travel time. Some information, however, has been made public. Delling et al. [5] note that Microsoft Bing's algorithm takes into account dozens of topological features such as the type of road, number of lanes, speed limit, and historical traffic data, and that the algorithm optimizes for simplicity as well by incorporating turn costs. Waze became infamous for especially accident-prone turns across traffic (an externality that likely arose as a result of optimizing more heavily than other commercial routing platforms on minimizing travel time) and has also since begun to explicitly optimize for simpler turns [39].

The introduction of the avoidance of "dangerous" areas in Rio de Janeiro by Waze [28] represents a large deviation from the fastest route paradigm. Waze also has incorporated avoidance of areas that cannot be entered legally by certain individuals, e.g. various settlements when driving in Israel [50]. Questions have been raised as to whether this type of safety routing merely enforces stereotypes and unfairly removes traffic (including potential retail customers) from poorer areas [28,39]. These concerns were also raised when a Microsoft patent that describes a means for helping pedestrians avoid areas where crime has been reported became public [15,33].

The Waze platform has also been accused of routing its users through many previously low-traffic neighborhoods. This has resulted in a number of externalities – including extensive frustration among residents (e.g. [18,29,57]) – and has led to calls for regulation of where mapping platforms can direct traffic [29]. Relatedly, a simulation study found that near-universal usage of fastest path routing during high-traffic times led to the redistribution traffic away from highways and onto more local roads [52]. Our research extends our understanding of traffic redistribution externalities to the large literature on alternative routing criteria and can help inform the policy debate about mapping platform regulation.

This research also builds on work that aims to provide some transparency to large-scale geographic systems that inform how we interact with the world, such as that by Soeller et al. [49], who developed a system for detecting personalization of political borders on Google Maps, and Chen et al. [4], who examined Uber surge pricing in San Francisco and Manhattan. This research takes a similar approach, but moves towards detecting the employment of alternative routing criteria (e.g. crime) rather than political borders or geographic biases in the sharing economy.

3 METHODS AND FRAMEWORK

In order to conduct a controlled evaluation of the externalities associated with alternative routing criteria, we needed three main components: the routing algorithms for each alternative criterion (i.e. beauty, safety, simplicity), a set of origin and destination pairs, and metrics to analyze the different externalities. For each origin-destination pair, and in aggregate across all origin-destination pairs for a city, we are then able to directly compare the routes generated by each algorithm. Below, we describe in detail our implementations of each of these components.

3.1 Alternative Routing Algorithms

We implemented three alternative routing approaches as well as a more traditional fastest-path algorithm to provide context when necessary. For our alternative approaches, we selected scenic, safety, and simplicity routing. As noted in Section 2.1, these three approaches have been validated by Golledge's work and have been the subject of substantial interest in the alternative routing algorithm literature. There is no consensus in this literature, however, on *how* to operationalize criteria like "scenicness" (which is referred to as "beauty" in some literature), safety, or simplicity in a routing algorithm. Additionally, there is also a lack of open implementations of these and other alternative routing approaches.

To address these issues, we developed our own framework that consists entirely of open-source software components and publicly-available data. We have released this framework for others to use and improve[2]. As noted above, the primary goals of the framework are (1) to make it easy for researchers and developers to consider externalities in their routing algorithm evaluations and (2) to provide a greater degree of standardization in routing algorithm evaluation more generally. Our framework is straightforward and has the benefit of being easily extensible to include additional alternative routing approaches and additional externality metrics not discussed in this paper (e.g. number of stop lights [14] or the diversity of neighborhoods along the route).

[2] https://github.com/joh12041/route-externalities

For implementation of our alternative routing and fastest-path algorithms, our framework utilizes the bidirectional Dijkstra implementation provided by the open-source GraphHopper Java library [3]. GraphHopper includes standard pathfinding algorithms and imports OpenStreetMap [4] road networks to build the underlying graph for routing. GraphHopper does not take traffic into account when determining the fastest path and instead bases travel time on road speed limits included in the OpenStreetMap data. All of our alternative routing approaches are included in our released framework.

3.1.1 Scenic Routing. Our implementation of scenic routing is designed to replicate the approach described by Quercia et al. [41], except where required by the nature of our study. Broadly, Quercia et al. collect the top k shortest paths between two points and then select the path that optimizes for beauty based on data derived from Flickr photo tags. We chose this approach because it is scalable to many geographic regions and complements open-source approaches as it relies on public social media data.

Quercia et al. base their underlying data on a LIWC-based [37] text analysis of the tags on geotagged Flickr photos. Specifically, they generate a 200m-by-200m grid across the city of interest in which each grid cell has a score based on its corresponding geotagged photos tags. They validated this approach with crowdsourced ground-truth beauty rankings. The notable variations from Quercia et al. in our system are as follows: our grid cells are slightly smaller and are not perfectly square (0.001 degree by 0.001 degree to necessarily speed up the algorithm) and we use Empath (Fast et al. [10], a validated open-source replacement for LIWC). In order to achieve better spatial coverage, we also add geotagged tweets [5] to the Flickr [54] tags that are used to score each grid cell. Despite these changes, as validation we note that we see similar trade-offs in travel time to those reported by Quercia et al.

In addition to defining scenicness, the Quercia et al. approach also needs a means of generating and ranking routes based on this alternative criterion. Quercia et al. use Eppstein's algorithm, which finds the k-shortest paths between an origin and destination. We do the same through GraphHopper, but find the k fastest paths where we cap k at 1,000 as Quercia et al. demonstrated diminishing returns with larger k values (we also examined k=10,000 and found similar results to those presented below, but with greater effect sizes). As is done in Quercia et al, each route is scored as the average beauty score of the grid cells through which it passes. The route with the highest average beauty score is selected and returned.

3.1.2 Safety Routing. We implement safety routing as closely as possible to descriptions of the technique used by Waze in Brazil. Though the specific details of the algorithm are not public, Waze notes that it avoids areas that have "higher-than-average homicide, car robbery, or drug trafficking rates" [28]. We focus our efforts with respect to safety routing on New York City [6] and San Francisco [7], both of which provide public crime data. We include data from all of 2016, retaining only the crimes that overlap with the categories mentioned above. It was also noted that Waze disregarded areas that had high numbers of drivers under the assumption that their users considered these areas to be safe [28]. Lacking this (private) data, we implemented a proxy: we disregarded highways (speed limit greater than 70 kilometers per hour) when avoiding roads in these areas.

Waze has not released the delineations of the areas in Rio de Janeiro that were designated "unsafe," so we tested several thresholds for determining which areas to instruct our algorithm to avoid. Waze chose 25 areas in Rio de Janeiro that are described as varying in size between a block and a neighborhood [28]. As such, we aggregate the crime data to census tracts, which in cities are generally of a size between a few blocks and a neighborhood. We then define a threshold for the average number of crimes (normalized by the area of the census tract) such that a certain percentage of census tracts (i.e. those above the threshold) are marked as "unsafe." We tested different variations of our threshold such that it removes 25%, 15%, 10%, 5%, or 1% of census tracts [8]. We report results for the 10% threshold, finding the results for the other thresholds to be very similar.

With a list of census tracts that exceeded the threshold for crime, we conducted safety routing in GraphHopper by using fastest path routing but avoiding all road segments that pass through these census tracts and do not have a speed limit greater than 70 kilometers per hour. Thus, a fastest path that does not pass through a blocked area will be unaffected while

[3] https://github.com/graphhopper/graphhopper
[4] https://www.openstreetmap.org
[5] https://dev.twitter.com/streaming/overview
[6] http://www.nyc.gov/html/nypd/html/analysis_and_planning/historical_nyc_crime_data.shtml
[7] https://data.sfgov.org/
[8] Actually operationalizing on all areas with "higher-than-average" crime would have blocked far too many census tracts – i.e. about a third of census tracts in each city.

a path that would have passed through a blocked area is rerouted to the fastest path that does not pass through a blocked area.

3.1.3 Simplicity Routing. We implemented simplicity routing as described by Shao et al. [47]. This approach scores the simplicity of a route as the sum of the complexity of each intersection through which it passes. The complexity of an intersection is modeled based on the degree of the intersection (i.e. number of intersecting roads) and what action is to be taken by the driver (i.e. going straight or turning). We re-use the *k*-shortest-paths framework from scenic routing, but instead of selecting the path with the highest average beauty, our simplicity algorithm selects the path that has the lowest complexity score (i.e. the simplest route).

3.2 External APIs

We also included routes from two external mapping platforms[9], Google and MapQuest, to address RQ3 (evidence of alternative criteria in routes from third-party platforms). We report results for routes that were gathered on weekdays from 5-7:30pm local time for both platforms, a time of high traffic. We also gathered routes from 2-4:30am local time (weekdays) for low traffic directions, but do not report these results as we found little difference in the actual routes (i.e. ~98% overlap in routes, just different expected travel times).

3.3 Origin-Destination Pairs

In order to robustly compare the routes provided by different routing optimizations, we needed a set of origin-destination coordinate pairs in each of our four cities. Ideally, analysis of routes would be done with a representative sample of route requests (e.g. from Google Maps or MapQuest server logs). However, this type of data is not publicly available. To address this issue, we take two approaches: (1) we adopt a common practice in the literature (e.g. [16,46,61]) and test the algorithms on randomly-generated origin-destination pairs from across a city's entire road network and (2) we also use publicly-available taxi trip datasets where available.

With respect to our randomly-generated dataset, we generate approximately 5000 origin-destination pairs each for San Francisco (California, USA), New York City (New York, USA), London (England), and Manila (Philippines). These cities were chosen to provide regional variation while still having sufficient English speakers to provide a good source of photo tags and tweets for our scenic routing algorithm (Empath is currently limited to English).

To provide additional context when possible, we verified the validity of these randomly-selected pairs by analyzing two datasets of actual route origins and destinations based on taxi pick-ups and drop-offs, one in San Francisco [38] and the other in New York City[10]. As we discuss below, for our route-level externalities (RQ1), we see the same high-level findings in our taxi-sampled and randomly-generated datasets, and so we only report the results for the randomly-generated pairs. For our community-level externalities (RQ2), we reach varying conclusions depending on whether we use the taxi-sampled or randomly-generated origin-destination pairs. As such, we focus our discussion of RQ2 on San Francisco and New York City and discuss both sets of results.

3.4 Externality Metrics

The first set of externalities that we examine are attributes of a route that are not traditionally considered in evaluations but that have been found to be important in how people choose routes (i.e. RQ1, *route-level* externalities). First, we evaluate the complexity of the route, which we measure in several ways: number of turns [14,26], number of left (or right in London) turns [24], and the metric that we use in simplicity routing [7], which takes into account the number of intersections passed through by a route and what action is taken at each intersection (i.e. turn or go straight). We also measure the beauty of all of the routes [1,14,24] – another desired property of routes as determined by Golledge – doing so in the same way as described above for our algorithmic implementation of scenic routing. Finally, due to the public outcry around Waze redirecting traffic from the highways into smaller neighborhoods, we also measured the percentage of time that each route was on highways (i.e. "motorways" in GraphHopper, which are operationalized as roads with speed limits greater than 70 kilometers per hour) and the percentage of time that each route was on slower neighborhood roads (i.e. streets below "secondary" in GraphHopper, as operationalized as roads with speed limits less than or equal to 40 kilometers per hour). For

[9] Waze does not provide a public API.
[10] www.nyc.gov/html/tlc/html/about/trip_record_data.shtml

all of these metrics, we compute 99% confidence intervals by bootstrap resampling the routes 1000 times (i.e. sampling routes with replacement from the approximately 5000 generated for each algorithm and city).

The second set of externalities relates to the *community-level* impact across all the routes considered (RQ2), with much of the motivation for these externalities arising from the public discourse around Waze [28,29], i.e. analysis of how traffic might be redistributed throughout a city and whether income appears to be a factor in this redistribution. We specifically focus on income because concerns have been raised that safety routing would also lead to the avoidance of poorer neighborhoods [15,28]. For these externalities, we focus on the cities in which we implemented safety routing (New York City and San Francisco, both of which also have detailed census data on income available[11]). We compute how income correlates with where an alternative routing algorithm redistributes traffic (as compared to traditional fastest-path algorithms). Specifically, for road segments that saw significantly increased traffic for a given alternative routing algorithm, we calculate the weighted average of household median income based on how much additional distance of roads passed through a given census tract as compared to the GraphHopper fastest algorithm. For example, if across all origin-destination pairs, there was an additional 8 km of routes in a census tract with a household median income of $60,000 and 2 km of routes in a census tract with a household median income of $50,000, then the weighted average would be $58,000 for roads that saw increased traffic. We do the same then for road segments that saw significantly less traffic and compare. We again compute 99% confidence intervals through bootstrap resampling with 1000 iterations on the origin-destination pairs.

3.5 Calculating Metrics for Commercial Mapping Platforms

The MapQuest and Google APIs provide the points that comprise the route that they return (i.e. latitude, longitude of enough points to accurately convey the geometry of the route). From these points, we can easily calculate the beauty of the MapQuest and Google routes through the same beauty grid-cell framework as used with GraphHopper. However, calculating the simplicity for these routes is less straightforward because the details of each intersection are not provided by Google and MapQuest. To overcome this problem, we use the map matching process developed by Newson and Krumm [34], which has also been implemented in GraphHopper[12]. This process converts the points into a corresponding path on the GraphHopper road network, from which simplicity can then be calculated as before. The Newsom and Krumm approach is not perfectly accurate, however, and so we enforce that the resulting matched route must be within 5% of the length of the original route in order to be considered in further analyses. Recall is generally around 85%, with the exception of Manila at 63%, which likely reflects differences in the underlying road networks of OpenStreetMap and the commercial mapping platforms.

4 RESULTS

In this section, we analyze and compare the routes (i.e. Google Fastest, MapQuest Fastest, GraphHopper Fastest, GraphHopper Scenic, GraphHopper Safe, GraphHopper Simple) according to the metrics described in the prior section.

4.1 RQ1: Route-Level Externalities for Alternative Routing Approaches

The routes corresponding to the San Francisco origin-destination pair featured in Figure 1 are illustrative of the type of route-level externalities that are seen between the different optimizations. In general, we see that the Google, MapQuest, and GraphHopper Simple paths share many of the same characteristics and are longer than the GraphHopper Fastest path, making more extensive use of highways. The GraphHopper Scenic route looks quite different from both the simplest and fastest routes, taking a more complicated path that passes through several popular areas such as Union Square before arriving at the destination. The GraphHopper Safe path also deviates substantially from the fastest path in order to circumvent the Tenderloin, an area of higher crime in San Francisco.

These trends in the route-level externalities that arise as a result of different optimization criteria can be seen in Figure 2, which shows the results of each route-level evaluation metric by Euclidean distance between the origin and destination (x-axes) and city (columns). We walk through Figure 2 in the sub-sections below and supplement the trends with statistics from Table 5 in the Appendix, which contains the actual values for each city and algorithm when the Euclidean distance is between 10 and 11 kilometers (i.e. one slice of the data in Figure 2). Of note, we present the externalities both normalized to

[11] https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml
[12] https://github.com/graphhopper/map-matching

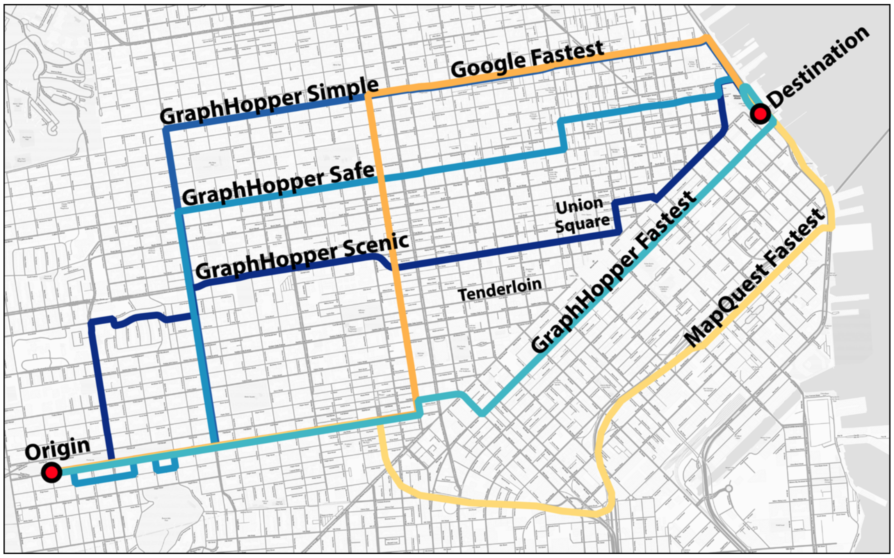

Figure 1. Routes given by each routing algorithm for an example origin-destination pair in San Francisco.

the natural baseline (e.g. GraphHopper Simple for simplicity measures) and also in absolute units when the units are readily interpretable (e.g. # of turns).

4.1.1 Route Complexity. Optimizing on beauty or safety substantially increases the complexity of routes, which has implications for driver safety and usability. As can be seen in rows 1 and 2 of Figure 2, the average number of turns and therefore complexity of a route increases significantly when comparing the fastest path to either the scenic or safer paths. Specifically, for route distances of 10-11 kilometers, the scenic path takes an additional 4-5 turns over the fastest path and the safer path takes on average an additional turn. Similar trends hold for the number of left turns (right turns in London) on a route (not shown) and the simplicity of a route (not shown, as measured by the number of intersections and action taken at each intersection). The simplest path is then an additional 2-3 turns shorter than the fastest path. For the safest path, this increase in complexity is the result of the added distance to circumvent a given area, but, for scenic routes, there are also significantly more steps per kilometer as well (third row of Figure 2).

4.1.2 Beauty. The scenic route is about 2-3x more beautiful (row 4 of Figure 2) than the other routes produced with other optimization criteria depending on the city and Euclidean distance. Notably, neither simplicity nor safety seems to correlate significantly with increased or decreased beauty.

4.1.3 Time on Highways and in Neighborhoods. Given concerns about the shift of vehicles away from highways and onto smaller neighborhood roads, rows 5 and 6 in Figure 2 consider the time spent by each route on each type of road. Scenic routes spend proportionally less time on the highway than the Google, MapQuest, or GraphHopper Fastest routes while GraphHopper Simple routes spend proportionally more time on the highway. Intuitively, this makes sense – many of the scenic spots in cities are not next to highways and taking a highway generally limits the number of intersections encountered. The magnitude of the differences varies across cities. On the low end, the scenic routes in London spend about 0.5% less of their time on the highway than the fastest path. At the high end, in San Francisco, the scenic routes spend about 9% less of the route on the highway than the fastest path, which corresponds to a 70% relative decrease in the amount of time spent on highways.

Scenic routes spend a significantly greater proportion of their travel time on slower roads, i.e. smaller roads that generally go through residential neighborhoods, foreshadowing their community-level effects highlighted below. The specific proportion of time spent on these roads varies greatly by city, but the GraphHopper Fastest, Simple, and Safe routes on average spend similar proportions of time on these roads whereas scenic routes tend to spend 25-50% (relative) more of their travel time on these roads.

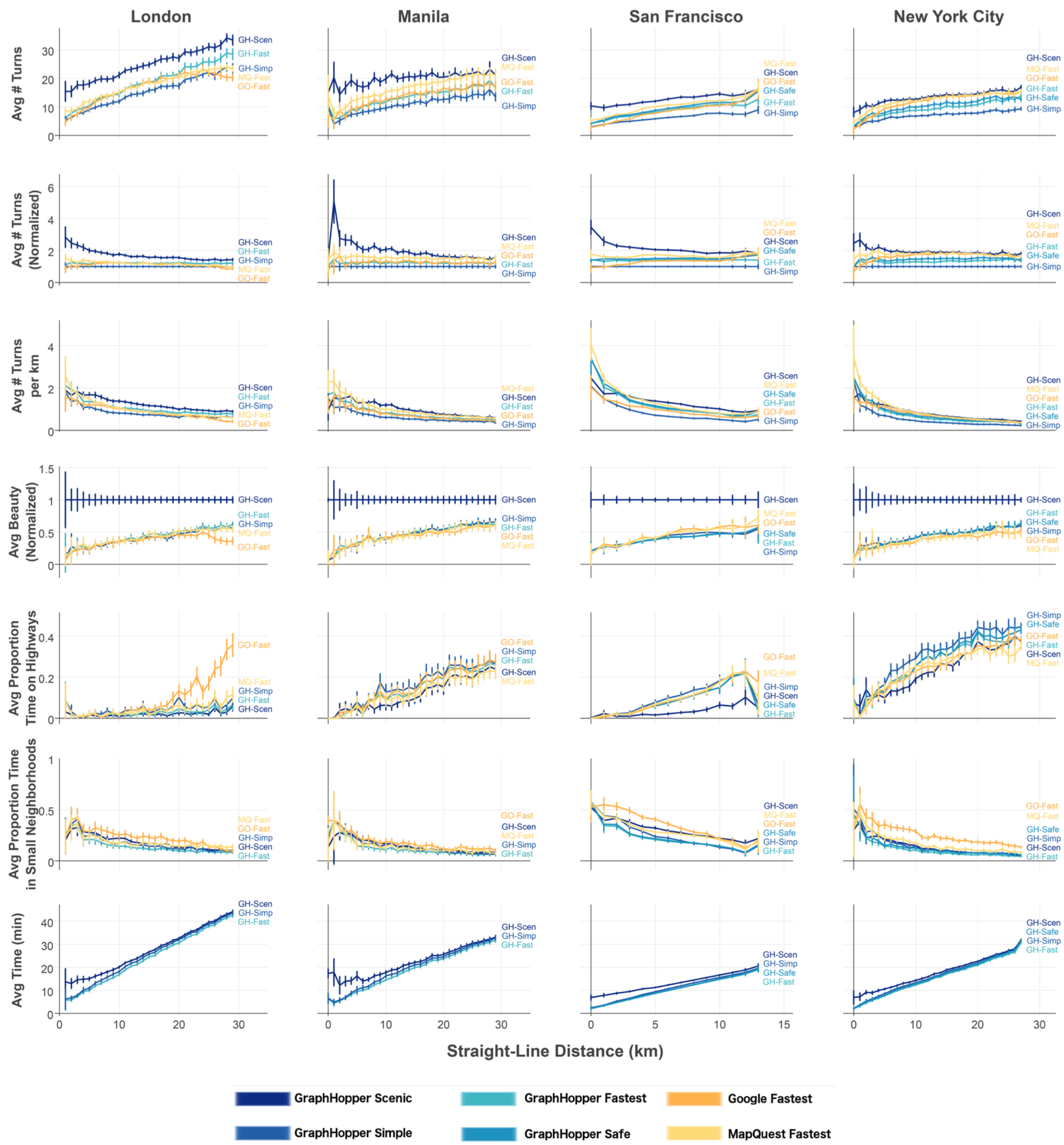

Figure 2. Comparison of the route-level externalities (rows) by city (columns), distance between the origin and destination pair (x-axis), and routing algorithm (lines). 99% confidence intervals calculated through bootstrap resampling. Average travel time (bottom row) for Google Fastest and MapQuest Fastest routes are not shown because any differences between them and the GraphHopper routes likely arise from variation in the underlying data for calculating travel time.

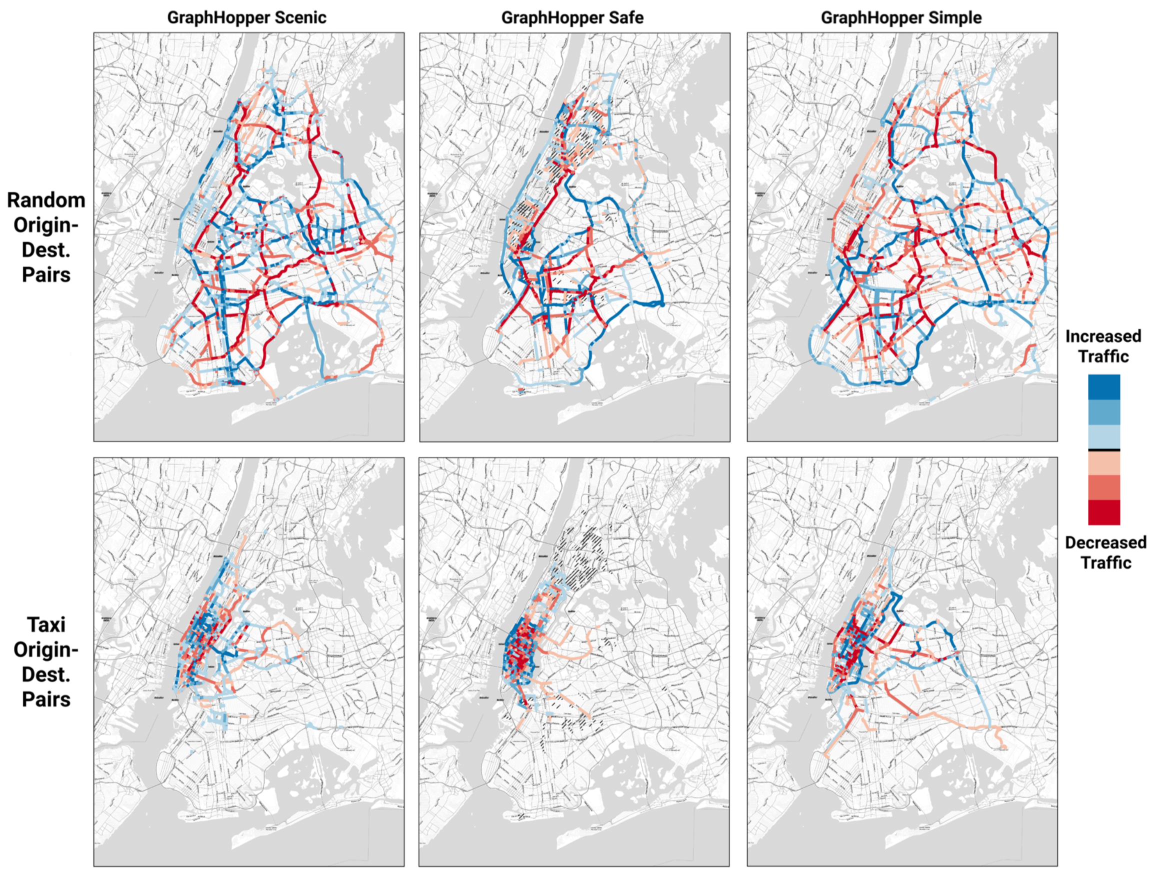

Figure 3. Comparison of alternative routing algorithms and GraphHopper Fastest algorithm, showing road segments with a significant change in the number of routes that pass over them across all origin-destination pairs for GraphHopper Scenic (left), Safe (middle), and Simple (right) routing in New York City. Results for random origin-destination pairs are shown on top and taxi-sampled origin-destinations are shown on bottom. Blue road segments had more routes pass over them with the alternative routing algorithm as compared to the GraphHopper Fastest algorithm, and red road segments had fewer routes pass over them with the alternative routing algorithm as compared to GraphHopper Fastest. The specific color thresholds were set by quantiles with the constraint mentioned above that blue represents increased traffic and red represents decreased traffic. Darker colors indicate a greater magnitude in the difference in number of routes passing over a given road segment. Black slanted lines in the GraphHopper Safe maps indicate blocked areas in safety routing.

4.2 RQ2: Community-Level Externalities of Alternative Routing Approaches

Motivated by public concern around the redistribution of traffic and disproportionate impacts on poor (or wealthy) areas, we also examined community-level externalities that may arise from optimizing on alternative criteria. Examples of these externalities are visualized in Figures 3 and 4, which show the roads that see significantly more or less traffic in New York City and San Francisco for scenic, safety, and simplicity routing. We show both the results from the randomly-generated origin-destination pairs, which provide good coverage of the entire areas, and the taxi-sampled origin-destination pairs, which are more representative of actual route concentrations.

4.2.1 Distribution of Traffic. For scenic routing, areas around parks see greatly increased traffic, as do popular tourist destinations and commercial districts. As can be seen in Figure 4, in San Francisco, the largest increases (>100 additional

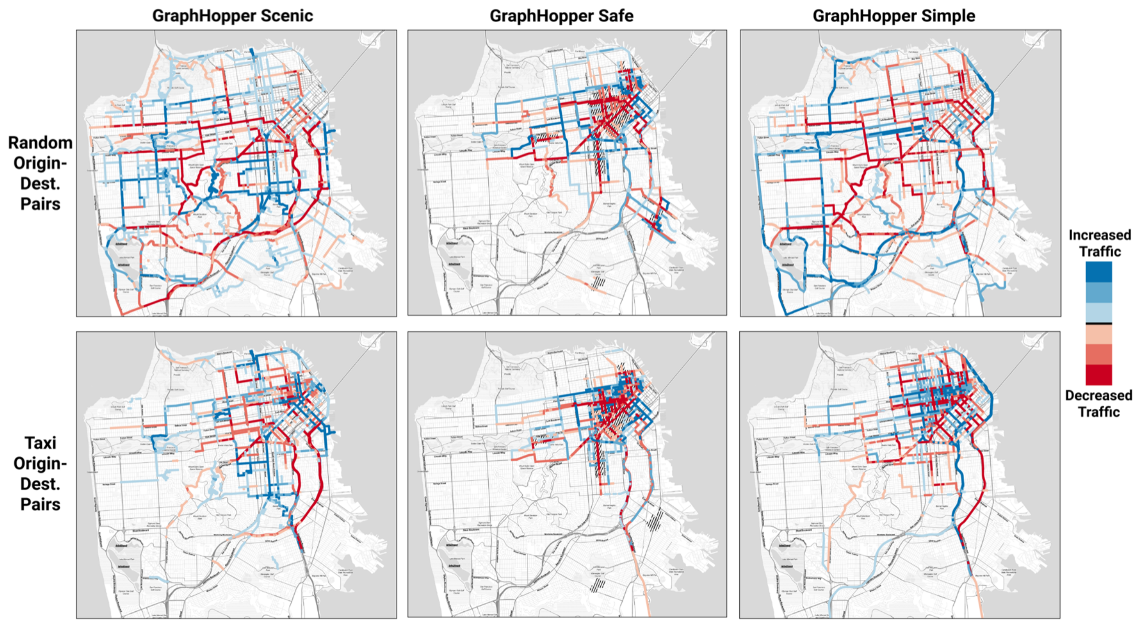

Figure 4. Comparison of alternative routing algorithms and GraphHopper Fastest algorithm, showing road segments with a significant change in the number of routes that pass over them across all origin-destination pairs for GraphHopper Scenic (left), Safe (middle), and Simple (right) routing in San Francisco. Results for random origin-destination pairs are shown on top and taxi-sampled origin-destinations are shown on bottom. Blue road segments had more routes pass over them with the alternative routing algorithm as compared to the GraphHopper Fastest algorithm, and red road segments had fewer routes pass over them with the alternative routing algorithm as compared to GraphHopper Fastest. The specific color thresholds were set by quantiles with the constraint mentioned above that blue represents increased traffic and red represents decreased traffic. Darker colors indicate a greater magnitude in the difference in number of routes passing over a given road segment. Black slanted lines in the GraphHopper Safe maps indicate blocked areas in safety routing.

routes out of the approximately 5000 analyzed in the random origin-destination pairs) occur around Golden Gate Park, the Embarcadero (popular waterfront region), Glen Park, and Mission Street as it passes through the Mission District (popular commercial district). The roads that see corresponding drops in traffic are often nearby highways or similarly large thoroughfares, which have high speed limits but are not always scenic. In New York City (Figure 3), scenic routing led to the largest increases in traffic (again >100 out of the approximately 5000 routes in the random origin-destination pairs) on roads that border the rivers, the roads around Central Park, and in Williamsburg (a rapidly gentrifying neighborhood in Brooklyn). Again, it is largely highways and nearby thoroughfares where the greatest decreases in traffic are seen.

The changes in traffic related to safety routing are much more localized, with increased traffic in order to circumvent the blocked census tracts being redirected to highways as well as more local roads that are immediately outside of the blocked areas. In San Francisco, the region that sees the largest decrease in traffic is the Tenderloin (a poorer neighborhood very close to downtown). In New York City, the taxi-generated routes show that the bulk of the traffic redistribution would occur to avoid high-crime areas in Manhattan, though the randomly-generated pairs indicate that regions of Brooklyn and Harlem would see traffic shifted to the highways as well.

Simplicity routing leads to a large increase in the amount of traffic on highways, as they have fewer intersections, but does not appear to favor or avoid any specific areas.

4.2.2 Income of Neighborhoods. Given concerns about safety routing criteria avoiding poorer neighborhoods, we also computed the weighted average of the household median income for roads that saw significantly increased or decreased traffic. The results are provided in Table 2.

Table 2. Average household median income of road segments that see significantly increased or decreased traffic with each alternative optimization criteria. Using origin-destination pairs derived from taxi routes (i.e. reflective of actual travel patterns as opposed to randomly generated) often demonstrates a higher income disparity between the types of roads preferred or avoided by a given alternative optimization. City-specific differences are evident as well.

Origin-Dest. Pairs	Change in Traffic	Household Median Income [99% Confidence Interval]		
		Scenic	Safe	Simple
New York City (Random)	Increase	$56,885 [55,484-57,555]	$59,110 [59,076-59,379]	$61,881 [61,838-62,160]
	Decrease	$55,902 [55,745-56,233]	$60,338 [59,256-62,561]	$59,283 [58,992-59,961]
New York City (Taxi)	Increase	$91,737 [90,997-92,590]	$91,870 [91,485-92,505]	$77,527 [74,021-78,618]
	Decrease	$88,209 [87,377-89,607]	$98,779 [96,982-101,834]	$89,106 [87,877-91,635]
San Francisco (Random)	Increase	$93,579 [93,024-93,891]	$87,352 [86,844-87,498]	$94,203 [93,868-94,383]
	Decrease	$99,039 [98,505-100,387]	$72,566 [70,590-73,357]	$98,315 [98,002-98,858]
San Francisco (Taxi)	Increase	$97,639 [97,077-98,528]	$76,660 [74,690-77,841]	$87,688 [87,543-88,184]
	Decrease	$92,135 [91,147-94,439]	$59,279 [56,607-60,863]	$73,772 [70,242-75,430]

Across the different algorithms and cities, we see mixed but persuasive results that alternative optimization can lead to large and disparate externalities in the types of areas that receive increased or decreased traffic. Scenic routing favors wealthier areas in both cities. The taxi-sampled (and arguably therefore more representative of actual traffic patterns) origin-destination pairs indicate that traffic on average moves towards wealthier regions - i.e. the average household median income of areas that see increased traffic is significantly higher than that of areas that see decreased traffic. For instance, in San Francisco for the taxi-sampled origin-destination pairs, the household median income of road segments that saw increased traffic was $97,639 while it was only $92,135 for road segments that saw decreased traffic. The randomly-sampled pairs in New York City indicate no significant correlation with traffic changes and income, though scenic routing causes traffic to move to less wealthy areas in San Francisco when using the randomly-sampled pairs.

Safety routing results in mixed effects across the two cities. In San Francisco, for both the taxi-sampled and randomly-generated origin-destination pairs, we see that safety routing moves traffic towards wealthier areas. The average household median income of areas that see increased traffic is $15,000 higher than that of the areas that see decreased traffic. In New York City, we see smaller disparities in the income of areas where traffic is redistributed, but the safety routing approach actually seems to cause traffic to shift on average towards *less* wealthy areas. Adding to the robustness of these results, we note that we see the same trends if we look at alternative metrics such as the proportion of increased and decreased traffic in areas with a household median income below a given threshold, e.g. $40,000.

4.3 RQ3: Alternative Criteria in External Mapping Platforms

Our results suggest that Google and MapQuest are likely incorporating simplicity as an optimization criterion in addition to travel time (like Bing [5]), generating simpler routes than would be expected under a pure fastest-path approach. For instance, Figure 2 shows that both Google and MapQuest provide routes that are similar to the GraphHopper Fastest and GraphHopper Simple routes. Interestingly, we calculated the percentage overlap between each combination of the various GraphHopper and external platform routes and found that MapQuest and Google are most similar to each other but that the highest overlap between Google or MapQuest and the GraphHopper routes is with GraphHopper Simple and not with GraphHopper Fastest.

Importantly, however, we do not see evidence of Waze-style safety routing being applied in either platform, i.e. neither Google nor MapQuest appear to be excluding any neighborhoods from their routes. More generally, we also observe no major externalities relative to GraphHopper Fastest in either commercial platform, aside from an increase in simplicity. This can be seen in Figure 5, which shows the significant differences (at 99% confidence) in the number of routes that pass over a given road segment when comparing Google and MapQuest Fastest versus GraphHopper Fastest. While many roads show different levels of traffic, there is no clear geographic concentration in roads that are favored or avoided. Looking back at Figures 3 and 4, we see that scenic and safety routing resulted in areas in the city where many roads all saw an increase (e.g. a popular and pretty neighborhood in scenic routing) or a decrease in traffic (e.g. an "unsafe" area in safety routing). We do not see these patterns appear for Google or MapQuest in Figure 5.

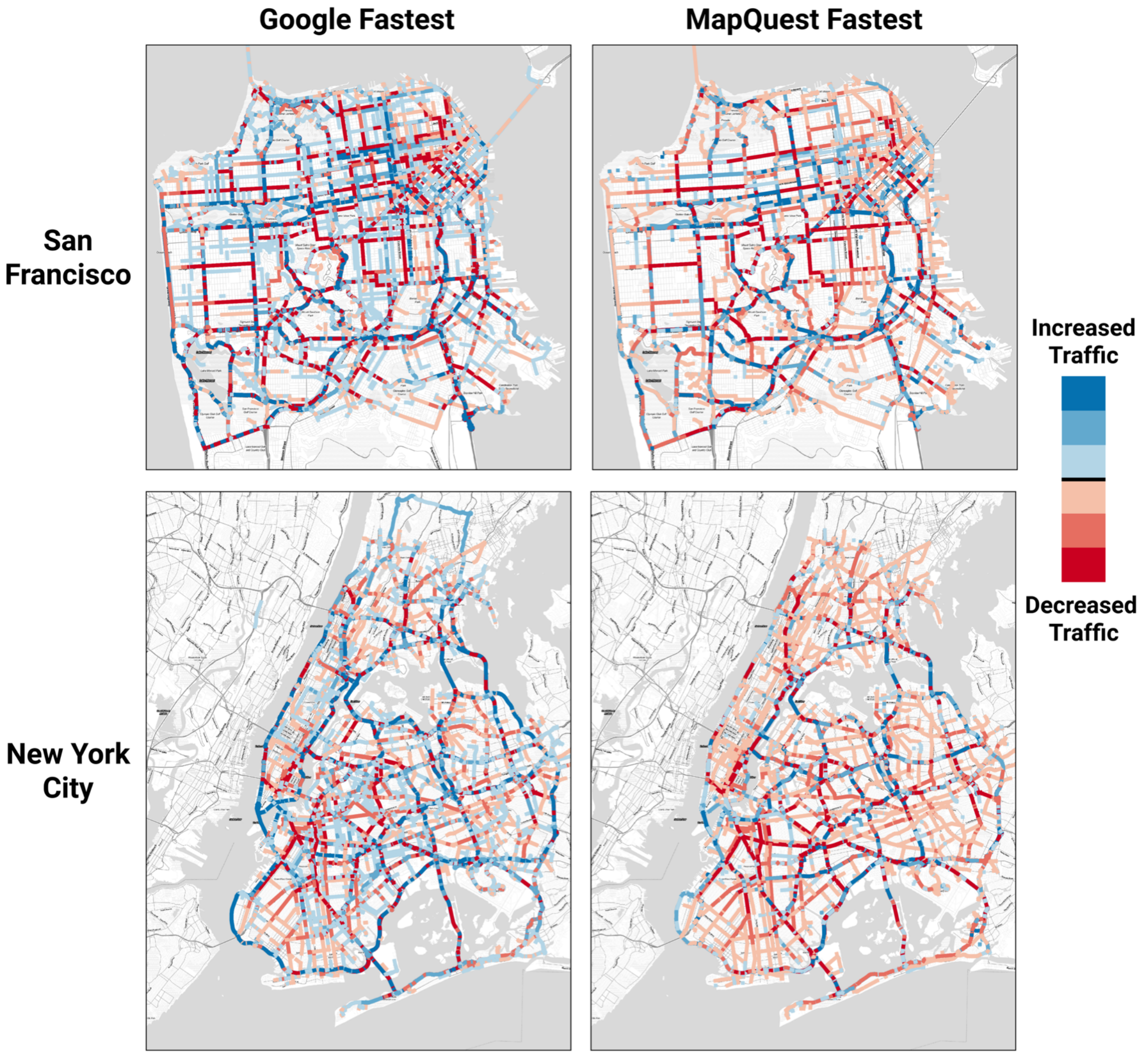

Figure 5. Comparison of third-party mapping platforms and GraphHopper Fastest algorithm, showing road segments with a significant change in the number of routes that pass over them across all origin-destination pairs for Google (left) and MapQuest (right) routing in San Francisco (top) and New York City (bottom). Results are shown for random origin-destination pairs. Blue road segments had more routes pass over them with the alternative routing algorithm as compared to the GraphHopper Fastest algorithm, and red road segments had fewer routes pass over them with the alternative routing algorithm as compared to GraphHopper Fastest. The specific color thresholds were set by quantiles with the constraint mentioned above that blue represents increased traffic and red represents decreased traffic. Darker colors indicate a greater magnitude in the difference in number of routes passing over a given road segment.

5 DISCUSSION

In this section, we discuss the implications of the above findings for both the design of routing algorithms and for public policy.

5.1 Societal Impacts of Alternative Routing

A clear high-level finding in the above results is that alternative optimization criteria are associated with important externalities that have not been previously considered. For instance, our results suggest that scenic routing (and safety routing to a lesser degree) led to substantially more complex routes involving several more turns on average. Turning, and specifically turns against traffic, are known predictors of collisions [30] and are less preferred by users [14]. Additional turns also present a usability challenge, with more complex routes leading to greater cognitive load [36], increased driver stress [56], and a greater likelihood of wrong turns and increased driving distance [16].

We also found that, if widely deployed, alternative optimization criteria such as beauty and safety would very likely lead to some of the externalities that have been a matter of public discourse and frustration with Waze [18,28,29,57]. Scenic routing redistributes traffic from highways into parks, popular areas, and onto slower, neighborhood roads. This raises concerns that optimizing on beauty could further contribute to the frustrations about increased traffic in previously low-trafficked neighborhoods [29]. These traffic increases on local roads have already led to calls by city council representatives in several cities to limit where mapping applications can direct their users [18]. Alongside the frustration and potential safety concerns of residents, high levels of traffic have also been tied to negative health outcomes [35].

By design, current safety routing approaches remove traffic from specific communities, which clearly could lead to economic and social impacts on those communities. While important to recognize, that this occurs in safety routing is not surprising. More surprising, however, is that traffic is not always just redistributed to surrounding (and likely similar) communities, but instead has far-reaching impacts and often is moved to highways and major thoroughfares that circumvent the larger area. Just as concerns have been raised about filter bubbles associated with the personalization of information consumption (e.g. [22,49]), safety routing may perform a similar function, allowing people to avoid areas that they do not want to see [28] and potentially shaping how we perceive the world [12]. A more balanced approach to safety routing might implement the avoidance of these areas only at times of low traffic and, during high-traffic times of the day, actually favor the "dangerous" areas so as to reduce the potential economic and social impacts of decreased traffic and visibility of these areas. Additionally, rather than focusing on high rates of crime, safety routing might instead focus on reducing the risk of collision by avoiding more dangerous driving maneuvers or crowded areas where accidents are more likely to occur.

In this paper, we explored previously-proposed alternative routing criteria with the concern that these could lead to adverse and disparate impacts in specific areas. One can also imagine, though, alternative routing criteria that lead to a greater diversity of experiences for the driver and more uniform impact on neighborhoods. We invite further discussion of what other alternative criteria might be considered that would arguably lead to *positive* externalities.

5.2 Towards Improved Routing Evaluations

Our results also suggest that traditional routing algorithm evaluations are insufficient to capture the potential for externalities. Specifically, the sole focus of traditional routing algorithm evaluations has been on potential increases in travel-time or distance, but this can hide important negative outcomes such as increased complexity (and its associated safety effects) and undesirable traffic patterns. Additionally, focusing just on travel-time and distance can lead to the conclusion that the trade-offs of an alternative optimization diminish rapidly with distance whereas we find externalities whose effects are relatively constant across distance. As can be seen in the final row of Figure 2, the increased travel time costs for the alternative optimizations diminish as the route distance increases (this matches what is seen by Quercia et al. [41] as well). However, this drop-off in magnitude of the trade-off is not nearly as immediate or does not occur when examining certain externalities, such as the number of turns or what types of roads comprise the routes.

5.3 Algorithmic Auditing

In this paper, we provided some of the first audits of routes generated by major mapping platforms. We did not find any evidence of major negative externalities associated with Google and MapQuest routes, with values for the externalities that we studied generally falling in the same range as those for GraphHopper Fastest and GraphHopper Simple.

As more commercial mapping platforms take steps like Waze has done to include notions of safety or other alternative optimizations in their algorithm, it is important that the research community continue to analyze the routes that these platforms are providing to ensure that any negative externalities are monitored. This is especially critical given the extent to which these platforms increasingly define the movement patterns of millions of people around the world. We have built our platform to accommodate the evaluation of routes from any external source by incorporating the map-matching component that converts a series of latitude-longitude coordinates to a route on the OpenStreetMap-based road network used internally by GraphHopper. Furthermore, the dataset of routes collected in the course of this research can serve as a baseline for future evaluations in order to detect changes in the routes provided by commercial mapping platforms.

5.4 Geography and Algorithms

In this research, we found that the externalities associated with a given routing algorithm varied across our four cities. For route-level externalities, we saw different effect sizes in each city, but the broad trends remained consistent. Among community-level externalities though, the nature of the trends changed from city to city. For instance, safety routing redistributed traffic to less wealthy areas in New York City and to more wealthy areas in San Francisco. The dependence of externalities on local geography extended to our choice of origin-destination pairs as well – i.e. randomly-generated vs. sampled from taxi routes. For route-level externalities, the choice of origin-destination pairs did not affect the trends, but the different sets of origin-destination pairs did lead to different conclusions for our examination of community-level externalities.

These results indicate that the interaction between routing algorithms and geography, especially when evaluating community-level effects, is highly dependent on the underlying urban structure and on origin-destination patterns within that structure. Care should be taken when generalizing results from one or two cities to other settings. Performance and results that vary across geographic contexts ("geographic human-computer interaction" [17]) has been highlighted in other domains as well – e.g., highly skewed performance of geolocation algorithms in urban versus rural areas [19], decreased precision predicting human perceptions of urban landscapes in a city when images and labels from a more distant city were used to train the model [32], different levels of anonymity in location-based social networks in different regions [43], and varying effectiveness of the sharing economy depending on the socioeconomic status of a given area [53]. Especially given the ubiquity of these algorithms (e.g. Google Maps alone has over a billion unique monthly users [6]), expanding alternative routing research to incorporate more geographic contexts will be important for guiding the design of these algorithms and supporting continued public discourse. We hope that our evaluation platform can assist in this endeavor.

6 FUTURE WORK AND LIMITATIONS

One of the large questions raised by this work is how might we design alternative routing algorithms in such a way as to realize their promised benefits while reducing the associated negative externalities. The literature provides some hints that are worth exploring. For instance, the simplicity routing literature notes that hybrid, multi-criteria approaches (e.g. balancing how much weight is given to simplicity and how much is given to travel time) often greatly reduce the complexity of routes while incurring minimal time costs [16,47]. Further study could examine whether multi-criteria optimizations, or other approaches that might directly consider externalities as a cost, show promise for reducing externalities. While we mentioned one way in which safety routing might be implemented in a more balanced form in Section 5.1, further insight might be gleaned from the urban studies literature (e.g. Jane Jacobs).

While the public discourse around Waze inspired some of this work, Waze currently does not provide a public API. Future work might pursue alternative means of collecting Waze routes, in part as an automatic means of detecting whether Waze routes are avoiding new areas. Furthermore, future work could also take advantage of open-source traffic data (e.g. [31]) to better understand the behavior of commercial mapping platform routes and explore how alternative routing algorithms react to changes in traffic as well.

We chose four cities as our geographic context for this study and sought to include cities from around the world. However, as noted in the discussion of geography and algorithms in Section 5.4, it is very probable that different impacts would be seen in other geographic contexts, such as in new cities or in suburban and rural areas. An interesting line of work would involve categorizing different cities based on how these algorithms perform so as to build a better understanding of how to more effectively target areas for study. In other words, are there classes of urban structures in which routing algorithms tend to have similar externalities? Of course, repeating our research in suburban and rural areas is also an important direction of future work.

We view this research as a first step towards understanding the externalities associated with various routing criteria. We sought to design our alternative routing algorithms based on actively-developed open-source libraries (e.g. GraphHopper, map matching) or published and validated methods (e.g. Empath [10], the *k*-shortest path approach developed by Quercia et al. [41]). However, there are many parameters and possible approaches to alternative routing, which, if executed differently, might lead to different results. Similarly, we analyzed the routes provided by Google and MapQuest for an initial directions request, but it is possible that these routes would be updated as they are driven in order to take advantage of shortcuts off of the highways. Finally, we built on previously-published methods to generate our alternative routes, but incorporating in human assessments of the resultant routes would provide additional certainty that the routes would be perceived as more scenic or simpler or safer.

7 CONCLUSION

In this paper, we provide the first robust assessment of the externalities associated with different alternative optimization criteria in geographic vehicle routing. We show that these externalities are substantial and would not be detected by traditional routing evaluations. For instance, we find that scenic and safety routing lead to more complex routes – increasing accident risks and other negative effects – as well as substantially increased traffic in various communities. The community-level impacts vary across different cities, however, demonstrating that evaluation across multiple geographic contexts is necessary in order to understand the impact of alternative routing approaches and highlighting the complex relationship between geography and algorithms more generally. We do not find evidence of negatives externalities arising in Google and MapQuest but release our evaluation platform so as to support continued evaluation and monitoring of commercial routing platforms. Finally, we discuss how algorithm designers might better balance the benefits of alternative optimization criteria with the externalities that can arise through their use.

ACKNOWLEDGMENTS

The authors would like to thank our colleagues in GroupLens Research at the University of Minnesota for their help in conceiving this project, especially Loren Terveen. This research was supported in part by the National Science Foundation (CAREER IIS-1707296 and IIS-1707319) and a Lichtenberg Professorship from the Volkswagen Foundation.

A ADDITIONAL TABLES AND FIGURES

Table 5. Rankings of the optimization criteria as well as values and associated 99% confidence intervals for origin-destination pairs separated by a Euclidean distance of 10 kilometers (i.e. numbers shown by the lines in Figure 2 when straight-line distance = 10). Though the magnitudes vary city to city, the rankings are largely stable across cities. Abbreviations: GHSc = GraphHopper Scenic; GHSi = GraphHopper Simple; GHSa = GraphHopper Safe; GHF = GraphHopper Fastest; GF = Google Fastest; MF = MapQuest Fastest.

Metric	Value [99% Confidence Interval] for origin-destination pairs with a Euclidean-Distance of 10km							
	London		Manila		San Francisco		New York City	
# of Turns	GHSc	**20.0** [19.0-20.9]	GHSc	**18.8** [17.4-20.0]	GHSc	**14.0** [13.6-14.5]	GHSc	**12.7** [12.2-13.3]
	MF	**14.4** [13.6-15.3]	MF	**13.6** [12.7-14.7]	MF	**12.0** [11.6-12.5]	MF	**12.4** [11.9-12.8]
	GHF	**14.2** [13.2-15.2]	GF	**11.5** [10.3-12.7]	GHSa	**11.3** [10.9-11.8]	GF	**11.4** [10.9-12.2]
	GF	**14.2** [13.1-15.2]	GHF	**11.0** [10.1-12.1]	GHF	**10.6** [10.2-11.0]	GHSa	**10.1** [9.48-10.6]
	GHSi	**11.6** [10.9-12.3]	GHSi	**9.36** [8.54-10.2]	GF	**10.4** [10.0-10.8]	GHF	**8.80** [8.32-9.30]
					GHSi	**7.67** [7.36-7.96]	GHSi	**6.80** [6.47-7.12]
# of Turns per kilometer	GHSc	**1.37** [1.30-1.46]	GHSc	**1.18** [1.10-1.25]	GHSc	**1.07** [1.04-1.11]	GHSc	**0.92** [0.88-0.97]
	GHF	**1.12** [1.04-1.20]	MF	**0.98** [0.91-1.05]	MF	**0.88** [0.85-0.92]	MF	**0.91** [0.88-0.95]
	MF	**1.09** [1.03-1.15]	GHF	**0.81** [0.73-0.88]	GHSa	**0.86** [0.83-0.90]	GF	**0.85** [0.80-0.90]
	GF	**1.06** [0.99-1.13]	GF	**0.77** [0.69-0.85]	GHF	**0.82** [0.79-0.86]	GHSa	**0.79** [0.74-0.83]
	GHSi	**0.79** [0.84-0.89]	GHSi	**0.57** [0.63-0.69]	GF	**0.73** [0.70-0.76]	GHF	**0.70** [0.66-0.74]
					GHSi	**0.56** [0.54-0.59]	GHSi	**0.49** [0.46-0.51]
Simplicity (Normalized)	GHSc	**1.46** [1.41-1.52]	GHSc	**1.71** [1.60-1.81]	GHSc	**1.49** [1.46-1.52]	GHSc	**1.54** [1.49-1.58]
	GF	**1.17** [1.11-1.24]	GHF	**1.16** [1.09-1.25]	GHSa	**1.28** [1.23-1.32]	GHSa	**1.30** [1.25-1.36]
	GHF	**1.14** [1.09-1.20]	GF	**1.13** [1.05-1.21]	GHF	**1.24** [1.20-1.29]	GHF	**1.24** [1.19-1.29]
	MF	**1.07** [1.02-1.12]	MF	**1.12** [1.04-1.20]	GF	**1.08** [1.05-1.11]	GF	**1.20** [1.14-1.26]
	GHSi	**1.00** [0.95-1.04]	GHSi	**1.00** [0.94-1.07]	MF	**1.04** [1.01-1.07]	MF	**1.08** [1.04-1.13]
					GHSi	**1.00** [0.97-1.03]	GHSi	**1.00** [0.96-1.04]
Beauty (Normalized)	GHSc	**1.00** [0.94-1.06]	GHSc	**1.00** [0.95-1.05]	GHSc	**1.00** [0.97-1.04]	GHSc	**1.00** [0.95-1.05]
	MF	**0.36** [0.30-0.43]	GHF	**0.41** [0.37-0.46]	GF	**0.53** [0.49-0.58]	GHSi	**0.39** [0.35-0.44]
	GHF	**0.34** [0.30-0.39]	GHSi	**0.40** [0.35-0.45]	MF	**0.53** [0.49-0.58]	GHF	**0.39** [0.35-0.43]
	GHSi	**0.34** [0.29-0.39]	GF	**0.38** [0.33-0.44]	GHSi	**0.50** [0.46-0.54]	GHSa	**0.38** [0.34-0.42]
	GF	**0.32** [0.26-0.37]	MF	**0.38** [0.33-0.43]	GHSa	**0.46** [0.43-0.49]	MF	**0.34** [0.30-0.38]
					GHF	**0.45** [0.42-0.48]	GF	**0.31** [0.27-0.36]
% Time Spent on Roads > 45 mph	MF	**2.84** [1.24-4.74]	GHSi	**17.5** [13.1-22.5]	GHSi	**15.0** [13.0-17.1]	GHSi	**27.8** [24.7-31.0]
	GF	**2.76** [1.18-4.91]	GF	**13.5** [9.56-17.4]	MF	**14.0** [12.2-16.1]	MF	**23.6** [20.5-26.4]
	GHSi	**2.54** [1.18-4.32]	GHF	**11.9** [8.31-16.3]	GHF	**13.5** [11.7-15.4]	GF	**21.4** [18.8-24.5]
	GHF	**2.07** [0.74-3.73]	MF	**10.9** [7.70-14.6]	GF	**13.5** [11.3-15.6]	GHF	**20.5** [17.6-23.7]
	GHSc	**1.50** [0.63-2.51]	GHSc	**6.18** [4.08-8.54]	GHSa	**13.3** [11.1-15.5]	GHSa	**20.5** [17.5-23.6]
					GHSc	**4.24** [3.19-5.41]	GHSc	**15.3** [13.0-18.0]
% Time Spent on Roads < 25mph	GF	**26.7** [23.0-30.4]	GHSc	**17.3** [14.8-19.9]	GF	**26.7** [24.5-29.0]	GF	**30.6** [27.8-33.5]
	GHSc	**21.8** [19.2-24.8]	GF	**17.1** [14.3-20.1]	GHSc	**24.2** [22.6-25.7]	MF	**18.4** [16.6-20.4]
	MF	**17.3** [14.4-20.4]	GHSi	**14.1** [11.1-17.4]	MF	**23.3** [21.2-25.3]	GHSc	**15.7** [14.5-17.0]
	GHSi	**16.4** [13.6-19.2]	MF	**13.1** [10.6-16.0]	GHF	**16.1** [15.0-17.4]	GHSi	**12.9** [11.5-14.4]
	GHF	**14.8** [12.4-17.6]	GHF	**11.5** [9.23-13.9]	GHSa	**16.1** [14.9-17.4]	GHSa	**12.1** [10.9-13.4]
					GHSi	**16.0** [14.6-17.5]	GHF	**9.95** [8.89-11.1]

Table 6. Total number of user-generated contributions (i.e. photos, tweets) for each city used in determination of beauty scores for scenic routing. Flickr photos from YFCC100M and Twitter tweets collected from May-August 2015.

Dataset	London	Manila	San Francisco	New York City
Flickr [# photos]	1,320,553	29,530	776,790	1,127,440
Twitter [# tweets]	262,216	328,179	200,609	685,172

REFERENCES

[1] Majid Alivand, Hartwig Hochmair, and Sivaramakrishnan Srinivasan. 2015. Analyzing how travelers choose scenic routes using route choice models. *Computers, Environment and Urban Systems* 50: 41–52.

[2] Mike Ananny, Karrie Karahalios, Christian Sandvig, and Christo Wilson. 2015. Auditing Algorithms from the Outside: Methods and Implications. In *ICWSM*.

[3] Hannah Bast, Daniel Delling, Andrew Goldberg, Matthias Müller-Hannemann, Thomas Pajor, Peter Sanders, Dorothea Wagner, and Renato F. Werneck. 2015. Route planning in transportation networks. *arXiv preprint arXiv:1504.05140*.

[4] Le Chen, Alan Mislove, and Christo Wilson. 2015. Peeking Beneath the Hood of Uber. In *IMC*. https://doi.org/10.1145/2815675.2815681

[5] Daniel Delling, Andrew V. Goldberg, Moises Goldszmidt, John Krumm, Kunal Talwar, and Renato F. Werneck. 2015. Navigation made personal: inferring driving preferences from GPS traces. 1–9. https://doi.org/10.1145/2820783.2820808

[6] Jillian D'Onfro. 2015. Here are all the Google services with more than a billion users. *Business Insider*. Retrieved from http://www.businessinsider.com/google-services-with-1-billion-users-2015-10

[7] Matt Duckham and Lars Kulik. 2003. "Simplest" Paths: Automated Route Selection for Navigation. In *International Conference on Spatial Information Theory*, 169–185.

[8] Abdallah El Ali, Sicco NA Van Sas, and Frank Nack. 2013. Photographer paths: sequence alignment of geotagged photos for exploration-based route planning. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 985–994. Retrieved April 4, 2017 from http://dl.acm.org/citation.cfm?id=2441888

[9] Samuel Elsmore, Irwan Fario Subastian, Flora Dilys Salim, and Margaret Hamilton. 2014. VDIM: Vector-based Diffusion and Interpolation Matrix for Computing Region-based Crowdsourced Ratings: Towards Safe Route Selection for Human Navigation. In *MUM* (MUM '14), 212–215.

[10] Ethan Fast, Binbin Chen, and Michael S. Bernstein. 2016. Empath: Understanding Topic Signals in Large-Scale Text. 4647–4657. https://doi.org/10.1145/2858036.2858535

[11] Kaiqun Fu, Yen-Cheng Lu, and Chang-Tien Lu. 2014. TREADS: A Safe Route Recommender Using Social Media Mining and Text Summarization. In *SIGSPATIAL* (SIGSPATIAL '14), 557–560.

[12] Birgitta Gatersleben, Niamh Murtagh, and Emma White. 2013. Hoody, goody or buddy? How travel mode affects social perceptions in urban neighbourhoods. *Transportation research part F: traffic psychology and behaviour* 21: 219–230.

[13] Robert Geisberger, Michael N. Rice, Peter Sanders, and Vassilis J. Tsotras. 2012. Route planning with flexible edge restrictions. *Journal of Experimental Algorithmics* 17, 1: 1.1. https://doi.org/10.1145/2133803.2133805

[14] Reginald G. Golledge. 1995. Defining the criteria used in path selection. *University of California Transportation Center*.

[15] Kevin Hamilton, Karrie Karahalios, Christian Sandvig, and Cedric Langbort. 2014. The image of the algorithmic city: a research approach. *Ann Arbor* 1001: 48109.

[16] Shazia Haque, Lars Kulik, and Alexander Klippel. 2006. Algorithms for reliable navigation and wayfinding. In *International Conference on Spatial Cognition*, 308–326.

[17] Brent Hecht, Johannes Schöning, Muki Haklay, Licia Capra, Afra J. Mashhadi, Loren Terveen, and Mei-Po Kwan. 2013. Geographic human-computer interaction. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 3163. https://doi.org/10.1145/2468356.2479637

[18] Steve Hendrix. 2016. Traffic-weary homeowners and Waze are at war, again. Guess who's winning? *Washington Post*. Retrieved from https://www.washingtonpost.com/local/traffic-weary-homeowners-and-waze-are-at-war-again-guess-whos-winning/2016/06/05/c466df46-299d-11e6-b989-4e5479715b54_story.html?utm_term=.1ff02cceab51

[19] Isaac Johnson, Connor McMahon, Johannes Schöning, and Brent Hecht. The Effect of Population and "Structural" Biases on Social Media-based Algorithms–A Case Study in Geolocation Inference Across the Urban-Rural Spectrum.

[20] David Jurgens, Tyler Finethy, James McCorriston, Yi Tian Xu, and Derek Ruths. 2015. Geolocation prediction in twitter using social networks: A critical analysis and review of current practice. In *ICWSM*.

[21] Jaewoo Kim, Meeyoung Cha, and Thomas Sandholm. 2014. SocRoutes: safe routes based on tweet sentiments. 179–182. https://doi.org/10.1145/2567948.2577023

[22] Chloe Kliman-Silver, Aniko Hannak, David Lazer, Christo Wilson, and Alan Mislove. 2015. Location, Location, Location: The Impact of Geolocation on Web Search Personalization. 121–127. https://doi.org/10.1145/2815675.2815714

[23] Julia Letchner, John Krumm, and Eric Horvitz. 2006. Trip router with individualized preferences (TRIP): Incorporating personalization into route planning. In *AAAI*, 1795.

[24] Shi-Chiang Li and Yongqiang Wu. 2012. South Florida Internet Based Route Choice Survey.

[25] YiRu Li, Sarah George, Craig Apfelbeck, Abdeltawab M. Hendawi, David Hazel, Ankur Teredesai, and Mohamed Ali. 2014. Routing Service with Real World Severe Weather. In *SIGSPATIAL* (SIGSPATIAL '14), 585–588.

[26] Ed Manley, Tao Cheng, and Andy Emmonds. 2011. Understanding Route Choice using Agent-based Simulation. *Proceedings of the 11th international conference on GeoComputation*.

[27] E.J. Manley, J.D. Addison, and T. Cheng. 2015. Shortest path or anchor-based route choice: a large-scale empirical analysis of minicab routing in London. *Journal of Transport Geography* 43: 123–139. https://doi.org/10.1016/j.jtrangeo.2015.01.006

[28] Aarian Marshall. 2016. Crime Alerts Come to Brazilian Waze, Just in Time for the Olympics. *Wired*. Retrieved from https://www.wired.com/2016/07/crime-alerts-come-brazilian-waze-just-time-olympics/

[29] Meghan McCarty. 2016. The Road Less Traveled? Not Since Waze Came To Los Angeles. *NPR All Tech Considered*. Retrieved from http://www.npr.org/sections/alltechconsidered/2016/06/13/481871496/the-road-less-traveled-not-since-waze-came-to-los-angeles

[30] Matt McFarland. 2014. The case for almost never turning left while driving. *Washington Post.* Retrieved from https://www.washingtonpost.com/news/innovations/wp/2014/04/09/the-case-for-almost-never-turning-left-while-driving/?utm_term=.9c07efa54881
[31] Randy Meech. 2016. Open traffic data is the best traffic data. Retrieved from https://mapzen.com/blog/open-traffic-data/
[32] Nikhil Naik, Jade Philipoom, Ramesh Raskar, and César Hidalgo. 2014. Streetscore-predicting the perceived safety of one million streetscapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 779–785. Retrieved February 14, 2017 from http://www.cv-foundation.org/openaccess/content_cvpr_workshops_2014/W20/html/Naik_Streetscore_-_Predicting_2014_CVPR_paper.html
[33] Wassim G. Najm, John D. Smith, and David L. Smith. *Analysis of Crossing Path Crashes.* National Highway Traffic Safety Administration.
[34] Paul Newson and John Krumm. 2009. Hidden Markov map matching through noise and sparseness. In *SIGSPATIAL*, 336–343.
[35] Ester Orban, Kelsey McDonald, Robynne Sutcliffe, Barbara Hoffmann, Kateryna B. Fuks, Nico Dragano, Anja Viehmann, Raimund Erbel, Karl-Heinz Jöckel, Noreen Pundt, and Susanne Moebus. 2015. Residential Road Traffic Noise and High Depressive Symptoms after Five Years of Follow-up: Results from the Heinz Nixdorf Recall Study. *Environmental Health Perspectives* 124, 5.
[36] Christopher JD Patten, Albert Kircher, Joakim Östlund, Lena Nilsson, and Ola Svenson. 2006. Driver experience and cognitive workload in different traffic environments. *Accident Analysis & Prevention* 38, 5: 887–894.
[37] James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates* 71, 2001.
[38] Michal Piorkowski, Natasa Sarafijanovic-Djukic, and Matthias Grossglauser. 2009. *CRAWDAD dataset epfl/mobility (v. 2009-02-24).*
[39] Linda Poon. 2016. Waze Puts Safety Over Speed by Minimizing Left Turns. *CityLab.* Retrieved from http://www.citylab.com/tech/2016/06/waze-puts-safety-over-speed-by-minimizing-left-turns-in-la/487577/
[40] Maaret Posti, Johannes Schöning, and Jonna Häkkilä. 2014. Unexpected journeys with the HOBBIT: the design and evaluation of an asocial hiking app. 637–646. https://doi.org/10.1145/2598510.2598592
[41] Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello. 2014. The shortest path to happiness: recommending beautiful, quiet, and happy routes in the city. In *HT*, 116–125.
[42] Daniele Quercia, Rossano Schifanella, Luca Maria Aiello, and Kate McLean. 2015. Smelly maps: the digital life of urban smellscapes. *ICWSM.*
[43] Luca Rossi, Matthew J. Williams, Christoph Stich, and Mirco Musolesi. 2015. Privacy and the City: User Identification and Location Semantics in Location-Based Social Networks. *ICWSM.*
[44] Nina Runge, Pavel Samsonov, Donald Degraen, and Johannes Schöning. 2016. No more Autobahn!: Scenic Route Generation Using Googles Street View. In *IUI*, 147–151.
[45] Sumit Shah, Fenye Bao, Chang-Tien Lu, and Ing-Ray Chen. 2011. CROWDSAFE: Crowd Sourcing of Crime Incidents and Safe Routing on Mobile Devices. In *SIGSPATIAL* (GIS '11), 521–524.
[46] Jie Shao, Lars Kulik, and Egemen Tanin. 2010. Easiest-to-reach neighbor search. In *SIGSPATIAL*, 360–369.
[47] Jie Shao, Lars Kulik, Egemen Tanin, and Long Guo. 2014. Travel Distance Versus Navigation Complexity: A Study on Different Spatial Queries on Road Networks. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (CIKM '14), 1791–1794.
[48] Monir H. Sharker, Hassan A. Karimi, and Janice C. Zgibor. 2012. Health-optimal Routing in Pedestrian Navigation Services. In *SIGSPATIAL* (HealthGIS '12), 1–10.
[49] Gary Soeller, Karrie Karahalios, Christian Sandvig, and Christo Wilson. 2016. MapWatch: Detecting and Monitoring International Border Personalization on Online Maps. 867–878. https://doi.org/10.1145/2872427.2883016
[50] Hunter Stuart. 2016. Waze Lets Israelis Avoid Palestinian Areas, but Not the Other Way Around. *Motherboard.* Retrieved from https://motherboard.vice.com/en_us/article/waze-lets-jewish-israelis-avoid-palestinian-areas-but-not-the-other-way-around
[51] Ivan J Tashev, Jeffrey D Couckuyt, Neil W Black, John C Krumm, Ruston Panabaker, and Michael Lewis Seltzer. 2012. *Pedestrian route production.* Google Patents.
[52] Jérôme Thai, Nicolas Laurent-Brouty, and Alexandre M. Bayen. 2016. Negative externalities of GPS-enabled routing applications: A game theoretical approach. In *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*, 595–601.
[53] Jacob Thebault-Spieker, Loren Terveen, and Brent Hecht. 2017. Towards a Geographic Understanding of the Sharing Economy: Systemic Biases in UberX and TaskRabbit. *ACM Transactions on Computer-Human Interaction.*
[54] Bart Thomee, David A. Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. 2016. YFCC100M: The New Data in Multimedia Research. *Commun. ACM* 59, 2: 64–73. https://doi.org/10.1145/2812802
[55] Martin Traunmueller, Ava Fatah gen. Schieck, Johannes Schöning, and Duncan P. Brumby. 2013. The Path is the Reward: Considering Social Networks to Contribute to the Pleasure of Urban Strolling. In *CHI EA* (CHI EA '13), 919–924. https://doi.org/10.1145/2468356.2468520
[56] Sudip Vhaduri, Amin Ali, Moushumi Sharmin, Karen Hovsepian, and Santosh Kumar. 2014. Estimating Drivers' Stress from GPS Traces. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 1–8.
[57] Naomi Zeveloff. 2016. Israelis Sue Waze Navigation App for Creating Neighborhood Traffic Jam. *Forward.* Retrieved from http://forward.com/news/israel/356487/israelis-sue-waze-navigation-app-for-creating-neighborhood-traffic-jam/
[58] Jianwei Zhang, Hiroshi Kawasaki, and Yukiko Kawai. 2008. A Tourist Route Search System Based on Web Information and the Visibility of Scenic Sights. In *Proceedings of the 2008 Second International Symposium on Universal Communication* (ISUC '08), 154–161. https://doi.org/10.1109/ISUC.2008.19
[59] Yan-Tao Zheng, Shuicheng Yan, Zheng-Jun Zha, Yiqun Li, Xiangdong Zhou, Tat-Seng Chua, and Ramesh Jain. 2013. GPSView: A Scenic Driving Route Planner. *ACM Trans. Multimedia Comput. Commun. Appl.* 9, 1: 3:1–3:18. https://doi.org/10.1145/2422956.2422959

[60] Brian D. Ziebart, Andrew L. Maas, Anind K. Dey, and J. Andrew Bagnell. 2008. Navigate Like a Cabbie: Probabilistic Reasoning from Observed Context-aware Behavior. In *UbiComp* (UbiComp '08), 322–331. https://doi.org/10.1145/1409635.1409678

[61] Dennis Zielstra and Hartwig Hochmair. 2012. Using free and proprietary data to compare shortest-path lengths for effective pedestrian routing in street networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2299: 41–47.

Received February 2017; revised April 2017; accepted June 2017