Proceedings of The National Conference
On Undergraduate Research (NCUR) 2019
Kennesaw State University
Kennesaw, Georgia
April 11-13, 2019

Identifying the Accuracy of GPS-Based Navigation Apps

Peter Hannon, Andrew Floyd, and Nicholas Alvino
Department of Computer Science
Kean University
1000 Morris Avenue
Union, New Jersey 07083 USA

Faculty Advisor: Dr. Daehan Kwak

Abstract

GPS-based real-time navigation apps via smartphones have become ubiquitous and are used in our everyday commutes. This has led to the exponential growth of real-time traffic data where it is utilized by several popular routing service providers such as Google Maps, Here, Bing Maps, Waze and MapQuest which all aim to provide the fastest route for travelers. Such map providers use historic and real-time crowdsourced traffic information to calculate the Estimated Time of Arrival (ETA) and construct the fastest route. However, there is no or limited information on whether these map providers are truly providing valid or accurate information. The purpose of this study is to gain insight and assess the quality of traffic information these popular online maps provide in terms of ETA. The methodology of this study includes two elements. The first element's objective is to collect traffic data from all the different map providers (mentioned above) by developing and deploying a web mining system to collect travel time estimates. A web mining system was also deployed to crawl the ground truth (or observed) traffic data from physical loop detectors and electronic toll tag readers that is publicly available provided by the New York DOT. For the second element of this study, a case study for the Manhattan, NY area was conducted with the observed and estimated traffic information to present a descriptive analysis. The data collected is grouped into three categories (rush hours, non-rush hours, and weekend). The ability to assess the quality of traffic information enables map service providers to improve the quality of their services and empowers the user to determine the services that best serves their needs.

Keywords: GPS, Accuracy, ETA, Routing, Navigation.

1. Introduction

GPS-based navigation apps are meant to conveniently direct a user from a given source to their destination. Many different apps with this functionality exist, namely Bing Maps, Google maps, Here, MapQuest, and Waze, which all report different Estimated Times of Arrival (ETAs). A real-time source of traffic data from a given sample area that could also be marked on navigation apps could provide insight into the accuracy of such apps, which could aid users in better determining ETAs and providing feedback on the accuracy of these navigation apps. By comparing the ETAs given by these navigation apps at a given time to open source real-time traffic data collected at the same time from the New York's Department of Transportation (NYDOT), a conclusion can be drawn to help determine the accuracy of these apps.

2. Methodology

2.1 Web Mining System

To compare the Estimated Travel Time (ETA) acquired from the five popular online maps, Bing Maps, Google maps, Here, MapQuest, and Waze, and the Actual Time of Arrival (ATA) collected from the New York's Department of Transportation (NYDOT) we built a system as shown in Figure 1. A dedicated server (System A) was programmed to crawl ETA data from the websites for Bing Maps, Google maps, Here, MapQuest, and Waze every 15 minutes. A separate dedicated system (System B) was programmed to crawl real-time traffic data from New York's Department of Transportation (NYDOT) every 15 minutes at the same time as System A across the same route using another python code.

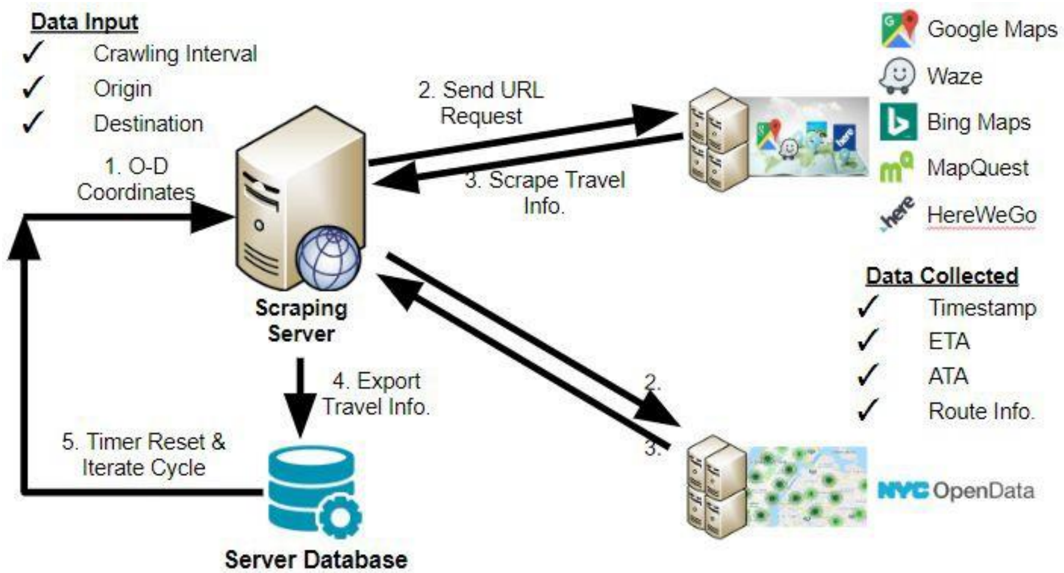


Figure 1. Layout of the Scraping System

2.2 Experiment Setup

An analyzation program calculates the average of all past ETA data collected by System A during the same weekday at the same time. Another interpretation program would calculate the average of the Actual Time of Arrival (ATA) collected by System B for the same weekday and at the same time. Figure 2 shows the NY Open data that is used to collect the ATA. The NY open data contains real-time traffic data on each road segment with a road ID. We identify the road segments (road IDs) that we collect through data parsing and cleansing.

Using these data, a reliable pattern of ETAs and ATAs could be drawn and compared. The data was then segmented into weekday (Monday-Friday) averages, and weekend (Saturday-Sunday) averages. Within that segmentation the data was further segmented into 8-hour intervals in order to clearly see data from morning rush hour, daytime traffic, and nighttime rush hour in order to see if these events have any effect on the data.

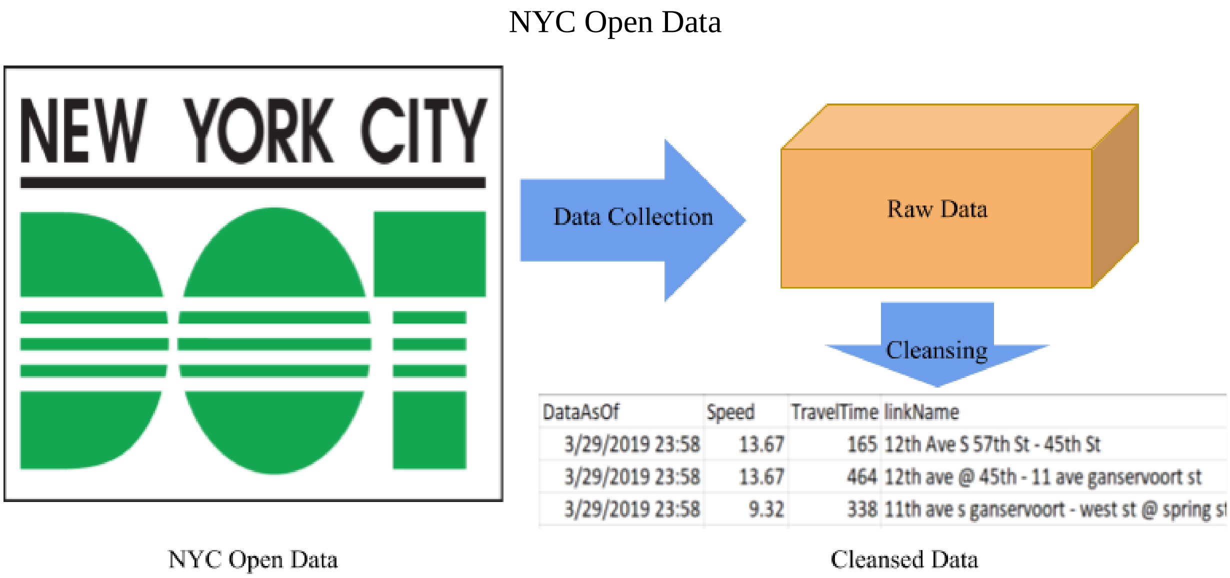


Figure 2. NY Manhattan Open Data

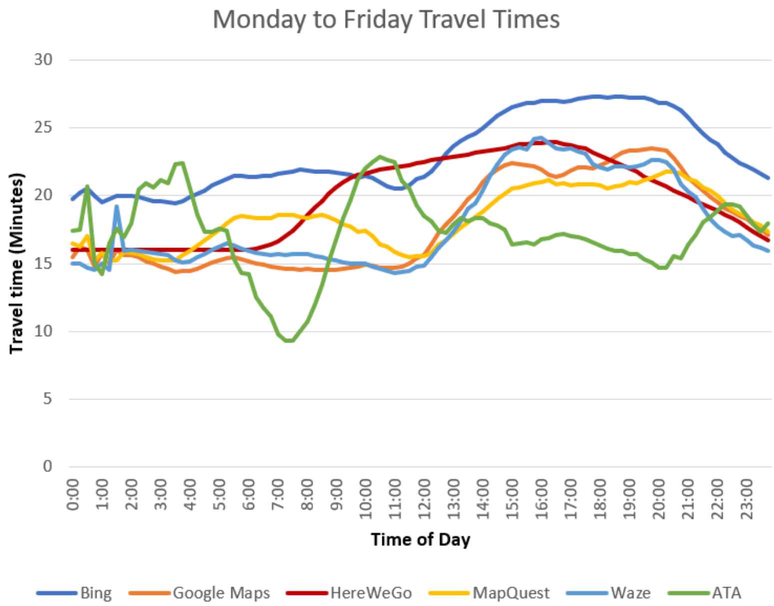
2.3 Case Studies on Routes

The case study on where the data was collected for the five popular online maps, Bing Maps, Google maps, Here, MapQuest, and Waze is shown in Figure 3. The starting point and destination were carefully selected so that all the applications showed a route that was identical and had a similar distance. In addition, each road segment was checked to ensure the ATA data collected from the New York Department of Transportation (NYDOT) was available. The selected start point is at 26 Morris Street, New York, Ny 10006, and destination at 830 12th avenue, New York, NY.

Figure 3. Case Study on identical Routes - NY

3. Results

The data that was collected over the course of 2 weeks. The graphs in figure 4 shows the ETA (Estimated Time of Arrival) for all mapping services that we compared, as the ATA (Actual Time of Arrival) for both weekdays (Monday - Friday), and weekend (Saturday - Sunday). This data shows an overview of how the arrival times differ. Any point in the graph where the ETA and ATA do not line up is an indicator of difference and therefore an incorrect ETA. Across both of these graphs in figure 4 there are few occasions where any of the graphed ETAs and the ATA meet at a common point, therefore the ETAs are mostly incorrect.



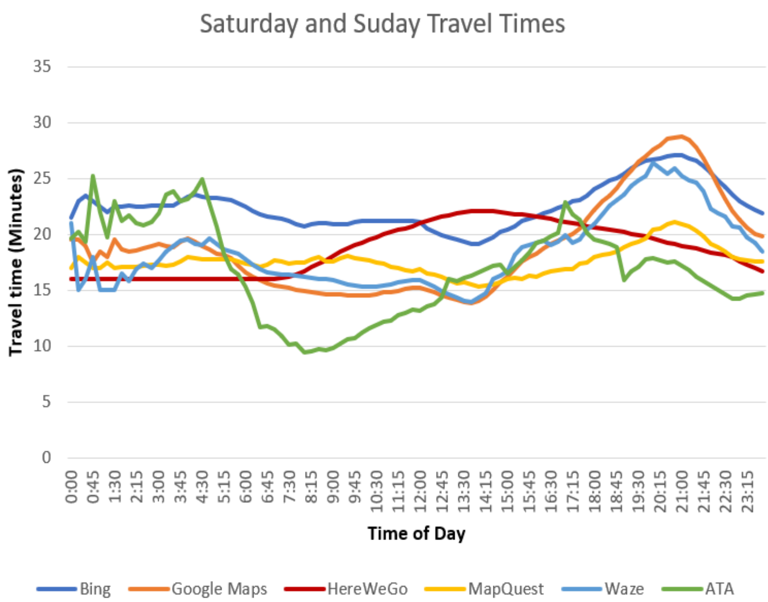


Figure 4. Collected travel time ETAs and ATA based on collected NYDOT data

Another set of graphs shown in figure 5 shows the difference between the ETA from all the mapping services and the ATA across the same weekday and weekend days. This provides a view of the data that is easier to interpret for the original purpose of this research, to find how severe the difference is and determine which service provides the most accurate estimation.

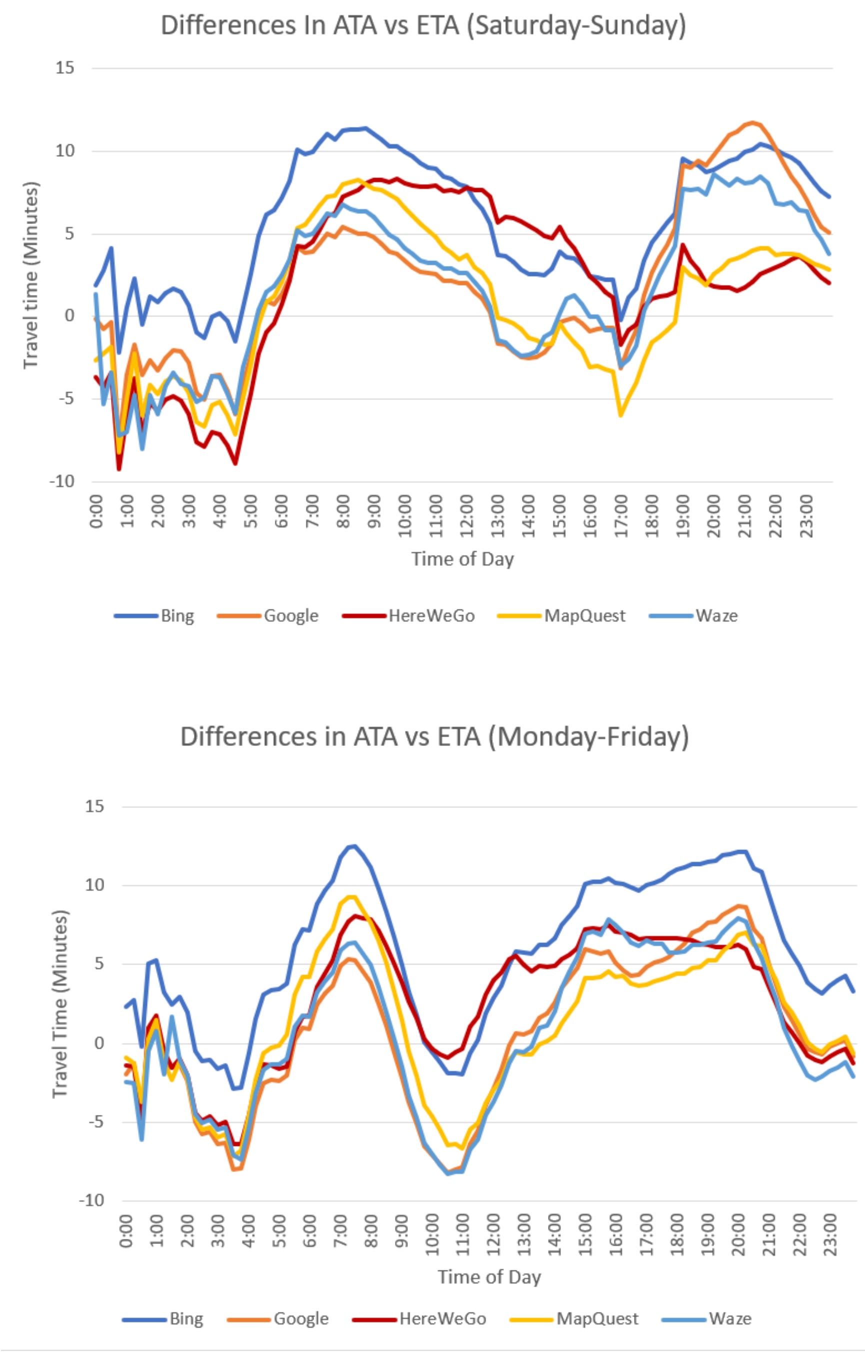


Figure 5. Difference between Average collected ETAs and the ATA for each 15-minute interval

Table 1 shows the interpretation of the data, by both segmenting the data across 8-hour intervals throughout the days from 0:00-8:00, 8:00-16:00, and 16:00-0:00, and showing the final average from the weekends or weekdays. These time slots were chosen as they accurately separate the data based on morning, midday, and nighttime traffic. The weekday data shows that on average MapQuest was the most accurate mapping service, having the least severe average difference between its ETA and the ATA, and Bing was the most inaccurate service, having the greatest average difference between its ETA and the ATA. The weekend data shows that on average MapQuest was the most accurate and Bing was the most inaccurate.

Table 1. Segmented numerical representation of collected data on ETAs most accurate to the ATA

	Time	0:00-8:00	8:00-16:00	16:00-0:00	Average
Weekday	BEST	Google	MapQuest	MapQuest	MapQuest
	WORST	Bing	Bing	Bing	Bing

	Time	0:00-8:00	8:00-16:00	16:00-0:00	Average
Weekend	BEST	Google	Google	HereWeGo	MapQuest
	WORST	HereWeGo	Bing	Bing	Bing

MapQuest appeared to be the most accurate service during the weekend despite never appearing as being the most accurate during any of the 8-hour time slots because it was on average the least incorrect throughout the day. On almost every time slot Bing Maps was the least accurate, so a conclusion can be easily drawn regarding how Bing Maps compares to the other mapping services. However, the Most accurate service during any of the given time slots appears to be mostly different and more spread out making it difficult to definitively assess which service is the most accurate. This data was only collected over 2 weeks and further collection may change these results and provide a clearer answer to the question of which service is most accurate.

4. Conclusion

GPS-based navigation systems are regarded as being a reliable source of navigation. By comparing the ETAs given by these navigation systems at a given time to open source real-time traffic data collected at the same time from the New York's Department of Transportation (NYDOT), a conclusion can be drawn to help determine the accuracy of these systems. These navigation systems all have different ways they calculate their routes and will therefore have different ETAs. Of the navigation systems tested (Bing Maps, Here, Google Maps, MapQuest, and Waze), the most accurate appears to be MapQuest, and the least accurate appears to be Bing Maps. This data could help the app developers to improve their systems for the benefit of their users, as well as equip users with the information on which system is able to boast the most accurate ETAs. In the future we intend to develop an intuitive website to collect and visualize the traffic data for future interdisciplinary research in Computer Science, Data Science, and Transportation. We also intend to keep collecting data and continue interpreting it in order to determine a clearer answer to the question of which mapping service is the most accurate.

5. Acknowledgements

Special acknowledgements and thanks to Daniel Henriques, Philip Kenny, and Dr. Daehan Kwak for their assistance and contribution to this work. This work is supported by Kean Students Partnering with Faculty (SpF) Summer Research Program.

6. References

1. Kwak, D., Kim, D., Liu, R., Nath, B., & Iftode, L. (2015). DoppelDriver: Counterfactual Actual Travel Times for Alternative Routes. Retrieved April 28, 2019.
2. Kwak, D., Liu, R., Kim, D., Nath, B., & Iftode, L. (2016). Seeing is believing: Sharing real-time visual traffic information via vehicular clouds. IEEE Access, 4, 3617-3631.
3. Kwak, D., Kim, D., Liu, R., Iftode, L., & Nath, B. (2014, November). Tweeting traffic image reports on the road. In 6th International Conference on Mobile Computing, Applications and Services (pp. 40-48). IEEE.
4. Sha, W., Kwak, D., Nath, B., & Iftode, L. (2013, February). Social vehicle navigation: integrating shared driving experience into vehicle navigation. In Proceedings of the 14th Workshop on Mobile Computing Systems and Applications (p. 16). ACM.
5. Ben-Elia and, Y. Shiftan, "Which road do I take? A learning-based model of route-choice behavior with real-time information", Transportation Research Part A: Policy and Practice, vol. 44, pp. 249-264, 2010.
6. E. Ben-Elia, R. Ishaq, Y. Shiftan, "If only I had taken the other road … ": Regret risk and reinforced learning in informed route-choice", Transportation, vol. 40, pp. 269-293, 2013.
7. Zhu S, Levinson D (2015) Do People Use the Shortest Path? An Empirical Test of Wardrop's First Principle. PLoS ONE 10(8): e0134322. doi: 10.1371/journal.pone.0134322
8. Bierlaire M, Chen J, Newman J. Modeling route choice behavior from smartphone GPS data. Report TRANSP-OR. 2010; 101016:2010.