Analysis of the Effect of Time of Day on Bus Speeds near a Bus Stop using GPS Data

I. Gökaşar[1], Y. Cetinel[1], A. A. Arısoy[1]

[1]Department of Civil Engineering, Boğaziçi University, İstanbul, Turkey

Abstract

Public buses are one of the most commonly used transit systems all across the globe. While they help in decreasing the number of vehicles in traffic, it is also important to minimize its impact on the overall traffic. For different times of the day, the traffic properties change and buses are influenced by different factors throughout the whole day. The objective in this study is to cluster the hours of the day accordingly, by using GPS data of the bus. By doing so, it will be possible to identify which properties of the bus stops have a significant relationship with the bus speeds in that period, respectively. The results will provide an understanding to determine for what conditions of traffic, the busses are influenced and in which way. Analyses show that during off-peak hours the bus speed is influenced the most by nearly every considered feature of the bus stop such as neighboring bus stop distances, the distance until there is a change in the number of lanes and the distance to an intersection of the road with a major road. However, between 16:00 and 17:00, where the traffic demand is high, none of the features inspected in this study for the selected bus stop, has an impact on the speed of the buses. Therefore, the outcomes may be interpreted as where the traffic demand decreases the features of the bus stop gain primacy in influencing the bus speed.

Keywords: *Bus trajectory data, GPS, bus speed, bus stop.*

1 Introduction

Public transportation is one of the key elements for a city's goal of achieving sustainable transportation system. Public busses have a significant importance for many major cities. In order to maximize the efficiency of this system, appropriate infrastructures, such as bus lanes and bus bays could be considered. However, failing to meet the requirements might cause ineffective system operation.

Bus stops are crucial elements for the efficiency of a successful public bus system. The characteristic properties of a bus stop, such as its location, availability of a bus bay at the stop will have an effect on both the efficiency of the transit system and the local traffic stream in the vicinity. Chien et al. (Chien and Qin, 2007) developed a mathematical model in order to improve the accessibility of the public bus system by optimizing the location and the number of stops on a segment of a road. For higher capacity buses, Fernandez (Fernandez, 1993) provided an expert system for the placement of bus stops used by these public vehicles. By optimizing bus stops, the number of buses in service can be reduced. Thus, operating costs can be minimized. A model by Saka (Saka, 2001) considers the fundamental relationship that exists among velocity, uniform acceleration or deceleration and displacement and among the average public bus operation speed, necessary fleet size and the potential system capacity. His model serves as an important decision tool for transit planners for determining adequate bus stop spacing in order to prevail network and traffic conditions. By adding more and more bus stops on the roads, nearby traffic experience unfavorable effects such as speed reduction and increased travel times. Koshy and Arasan (Koshy and Arasan, 2005), by constructing a model with mixed traffic, determined that the bus stops at curbside reduces general traffic speed more than 25%; therefore, requiring a bus bay built. Traffic lights, increase or decrease of number of lanes, number of lanes, intersection with a street or an avenue near any bus stop can be influencing factors for the impact of a bus stop on the nearby traffic. The behavior of the buses at the bus stops, combined with the properties of the bus stop, influence the traffic in some form.

Dwell time at the bus stops is a critical component for both the travel time of the busses and other vehicles in traffic. As dwelling duration increases due to alighting or boarding bus passengers, the negative impact on the overall traffic and other vehicles will also increase and it will damage the level of service quality of the public buses (Mazloumi et al., 2010) . Influencing factors on dwelling duration include passenger activity; lift operations, low floor bus, time of the day and route type (Dueker et al., 2004). The delays that the buses experience at the bus stops are linked explicitly in order to obtain a reliable algorithm that can be adopted to predict bus arrival times (Padmanaban, 2009). Min-Tang Li et al. (Li et al., 2006) developed a model, which estimates dwelling duration, considering the number of alighting and boarding passengers. A different model built by Milkovits (Milkovits et al., 2008) additionally considers other factors that has an impact on the dwell time such as crowding, fare type and bus design, using automatic passenger and fare and vehicle location data. In order to reduce the deviation from schedule due to dwelling duration, Turnquist (Turnquist, 1982) proposes to increase the bus stop spacing by removing some of the stops of a route and zone scheduling.

For analysis, using GPS data provided from the equipment found in the buses is the most popular method. Travelers can track the location of the bus that they want to use and operators can evaluate the level of service of the transit system by utilizing this information. Researchers make use of this type of data in their studies as well. Zeng (Lin and Zeng, 1999) made use of the GPS information to analyze the real-time bus arrival time prediction for interpreting the level of services whereas Stoll et al. (Stoll et al., 2016) used this data for visualizing and identifying congestion critical points in urban arterials. Pinelli et al. (Pinelli et al. 2013) utilized GPS dataset to develop a bus stop identification method. GPS information is also utilized for detecting travel modes such as walking, private vehicle, public bus, and subway and commuter rail in the study of Gong (Gong et al., 2012). Although there are multiple ways to make use of GPS data, there are also its difficulties for obtaining accurate results.

The GPS dataset on hand can be enormous and should be manipulated if necessary for the aimed objective. Stopher et al. (Stopher et al, 2005) developed useful methods in order to convert the data collected from GPS devices into a more manageable format. However, there are still difficulties that can occur during the analyses. Csiszár and Sándor (Milkovits et al., 2008) noted in their studies that constructing a database based on the tracking system installed on the vehicle and classifying movement and dwell phases was their research's most challenging objective. There is also the issue of inaccurate data, which can create significant issues if not fixed before the analysis. Even though theoretically challenging, combining GPS monitoring, accelerometry and GIS technologies seems to be promising for interpreting the transport-related physical activity (Oliver et al., 2010). One must be elaborate while using this type of data to avoid significant errors.

This study focuses on discovering a relationship between the speed pattern of the public buses by utilizing GPS data and the properties of a bus stop, which is stored in a data matrix. The properties of the bus stop include information such as the number of lanes where the bus station is, if there is a change in the number of lanes or a traffic light exists near the bus stop. The final step is to determine if there is a time dependency of the effects of geometrical properties on the speeds of the selected buses.

2 Methodology

For the analysis, Yıldız Technical University (Figure 1) bus stop is selected. The location of the bus stop is in Besiktaş and it is right after the Barbaros transition road which leads to the 15th of July Martyrs Bridge. There are 3 lanes where the bus stop is located and a bus bay for this bus stop. The distance between the previous (Barbaros Avenue bus stop) and the next (Ertuğrul Sitesi bus stop) bus stop is divided into sections of 20 meters and 56.785 data points (Figure 1) and it is analyzed.

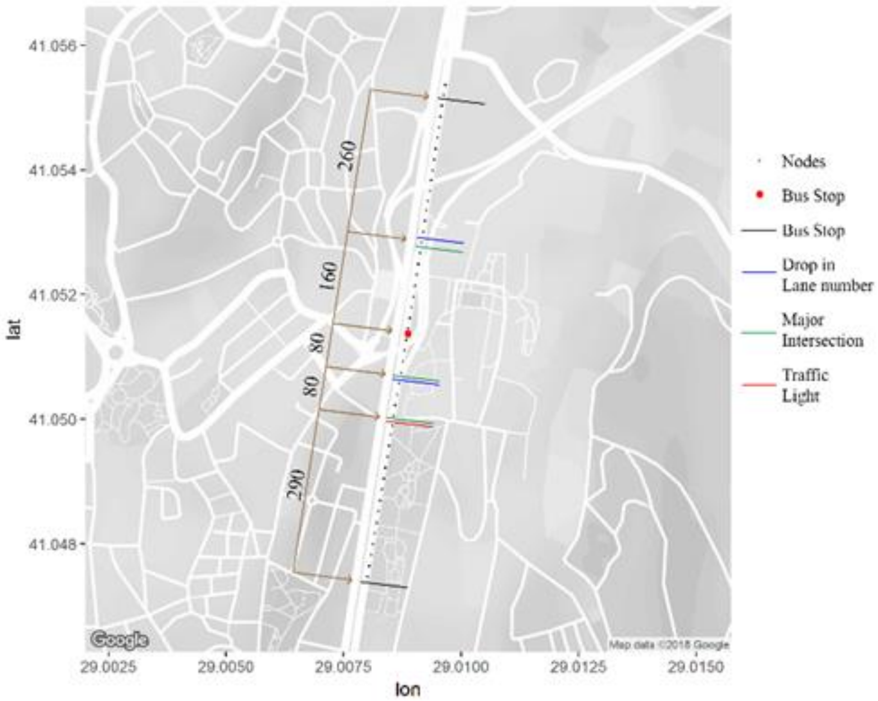

Figure 1. Position of the selected bus stop.

Using the GPS data of the bus routes (30A and U2) and the k-means algorithm, clusters are constructed by inspecting the speed and time stamp of the bus. The route information of the bus is collected from the IETT (Istanbul Electric Tramway and Tunnel Establishments). The data structure consists of 9 features; Geopoint ID, time, latitude, longitude, distance, time difference, time difference, speed, node ID and position (Table 1).

It is expected to have meaningful clusters where each cluster represents a specific period of the day. The objective here is to determine which geometrical property of the bus stop has a significant relationship with the speed of the buses for different time intervals, where clusters represent these time intervals. For achieving this objective linear regression model is used.

Table 1. Sample data structure.

GeopointId	Time	Latitude	Longitude	Distance (m)	Time Difference (sec)	Speed (kmh)	Node (id)	Position (m)
5379093	30/04/16 06:17	41.055614	29.009445	172.457	16	38.803	47	450.1
5379094	30/04/16 06:17	41.053276	29.008951	238.962	15	57.351	34	204.6
5379095	30/04/16 06:17	41.050720	29.008415	261.160	16	58.761	20	-68.9
5379096	30/04/16 06:18	41.049400	29.008200	133.739	16	30.091	12	-221.1
5379097	30/04/16 06:18	41.047320	29.007729	213.266	16	47.985	1	-438.3
5379185	30/04/16 06:41	41.048767	29.008158	85.726	15	20.574	8	-300.1

3 Results

Considering this specific bus stop (Yıldız Technical University), 5 clusters are constructed using the k-means algorithm (Figure 2). Each number in the cluster symbolizes the hour of the day (6 represents the time of day between 06:00 and 07:00). By observing the results, logical labelling of these clusters is possible.

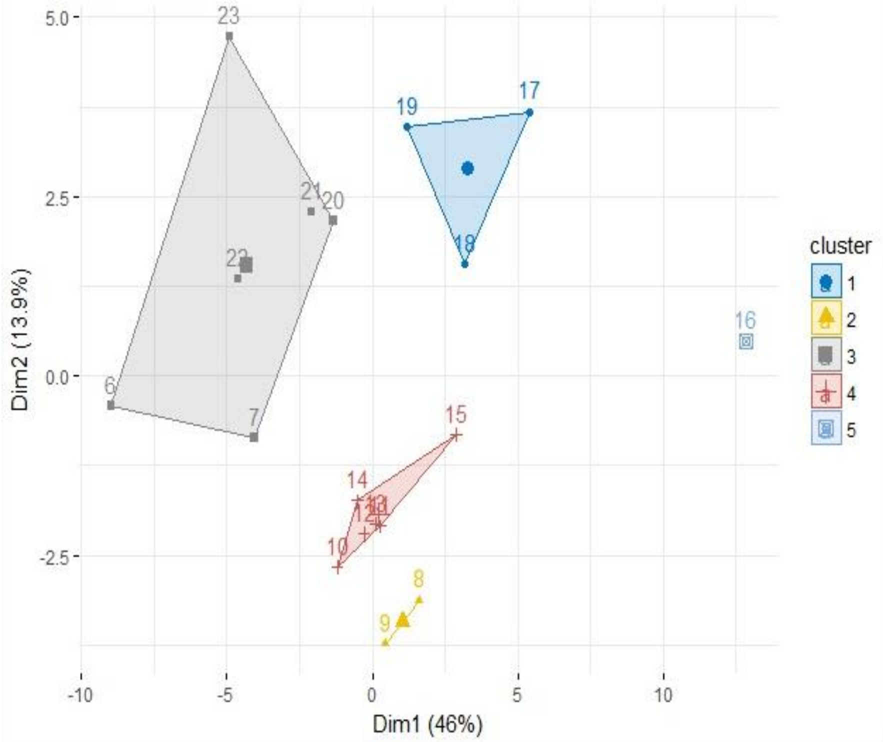

Figure 2. Clusters of bus speed patterns.

Cluster 1 (colored with blue) covers the evening peak hours. Cluster 2 represents the hours of the day where the traffic is stabilized but congested, which means that the new vehicles that enters the system is less than the previous cluster. Cluster 3 is formed due to properties of the location of the bus stop. Cluster 4 shows the off-peak hours and cluster 5 represents the critical hour where there is a large increase in traffic demand. In time interval between 16:00-17:00, travelers aim to start their trips to avoid the peak hour traffic afterwards. This is probably why this cluster has a different effect on the bus speed. Figure 3 shows the bus speeds for the mean of each cluster.

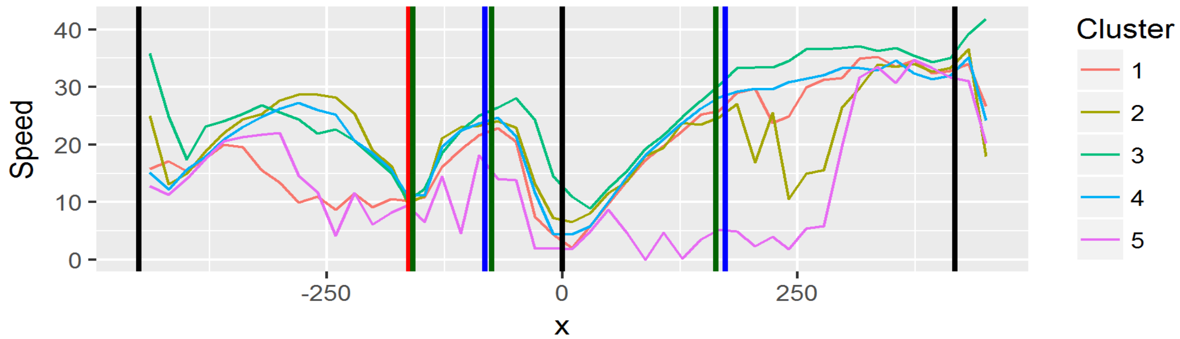

Figure 3. Speed patterns of clusters (black: bus stop, green: major intersection, blue: drop-in-lane-number, red: traffic light).

Table 2. Linear regression results.

Parameter	Estimate	Std. Error	t value	Pr(>|t|)	Significance Codes
(Intercept)	-0.84444	1.728448	-0.489	0.6256	
dist.stop	0.057533	0.012885	4.465	1.26E-05	***
dist.drop	-0.07581	0.034896	-2.172	0.0309	*
dist.light	0.014128	0.008221	1.719	0.0871	.
dist.major	0.128821	0.05044	2.554	0.0113	*
cluster.1	6.962239	1.37546	5.062	8.61E-07	***
cluster.2	8.608182	1.37546	6.258	1.93E-09	***
cluster.3	13.02619	1.37546	9.47	< 2e-16	***
cluster.4	9.717269	1.37546	7.065	1.97E-11	***

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The distance to the neighbor bus stops (dist.stop) is the most significant feature when the bus speed is considered. After that, decrease in the number of lanes (dist.drop) and intersection with a major road (avenue) are the next significant variables that are influencing the speed of the bus (Table 2). The least significant variable is the distance to the traffic lights (dist.light) from the bus stop, concluding that all available features of the bus stop has a significant relationship with the speed of the bus. After that, it is determined if there is a significant relationship of these features with clustered different periods of the day (Table 3).

Table 3. P-Values.

Cluster Index	(Intercept)	dist.stop	dist.drop	dist.light	dist.major
Cluster 1	0.015730360	0.043292120	0.217674630	0.013516010	0.300501370
Cluster 2	0.006975474	0.082143480	0.616892720	0.754086050	0.456873200
Cluster 3	0.000182500	0.004382246	0.139747770	0.113608110	0.107949510
Cluster 4	0.050460690	0.000008176	0.054275120	0.328501610	0.048113990
Cluster 5	0.459828079	0.565289500	0.471055550	0.804310920	0.224460230

For the cluster of evening peak hours (Cluster 1), only the traffic light and neighbor bus stop distances to this bus stop are the significant features. The major roads and lane drops are not significantly related to speed values of this period of the day.

For the time interval represented by the 2nd and 3rd clusters, only the neighbor bus stop distances have a significant influence. The other features do not have the influence that they had for the speed of the bus for the whole speed data.

The speed values of the off-peak hours (Cluster 4) are significantly related with neighboring bus stop distances, the distance until there is a lane drop and the distance to an intersection of the road with a major road.
The final cluster (Cluster 5) which only consists of one period has no significant relationship with any of the features. This might be due to the large demand increase for this period. The increased demand in this hour causes the buses to travel with consistently low speeds without any sudden changes, whether there is a traffic light or an intersection with a major road (Figure 4).

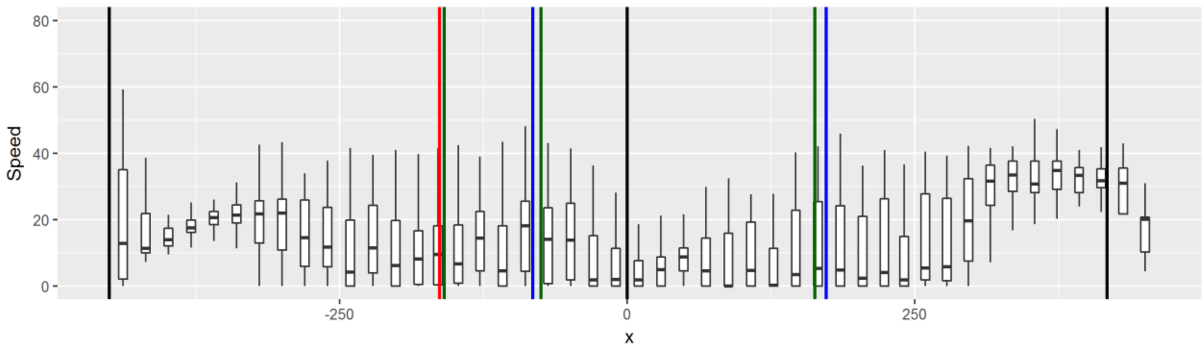

Figure 4. Speed patterns of buses between 16:00-17:00 (black: bus stop, green: major intersection, blue: drop-in-lane-number, red: traffic light).

In Table 4, how different features and different hour of the day influences the bus speeds is given as a summary.

Table 4. Coefficients of the separate models for each cluster.

Cluster Index	(Intercept)	dist.stop	dist.drop	dist.light	dist.major
Cluster 1	6.2114920	0.04436113	-0.07215239	0.035023936	0.08735702
Cluster 2	9.9630380	0.05395226	-0.04134895	0.006094533	0.08905113
Cluster 3	10.101064	0.06394468	-0.08654563	0.021888642	0.13650619
Cluster 4	4.8469980	0.10550430	-0.11134440	0.013097833	0.16547096
Cluster 5	2.9690930	0.01990067	-0.06763519	-0.005462654	0.16572016

4 Conclusion

In this study, it is aimed to analyze the relationship between the geometrical properties of a bus stop and bus speeds at different periods of the day. The geometrical properties of each bus stop can be distinct. Thus, a bus stop in the city center where there is a significant amount of traffic demand throughout the whole day is selected.

After the analyses, it was determined that for different periods different features gain importance in influencing the speed of the buses. Only in the period between 16:00 and 17:00, it was observed that no feature is significant, which was probably due to high traffic demand. During off-peak hours, with the decreased amount of traffic demand, bus speeds are mainly influenced by the geometrical properties nearby. Due to these findings, it may be concluded that as the traffic demand decreases, the geometrical properties remain as the main factors that influence the bus speeds.

This methodology can be easily implemented in cities, which have very large bus network. Also, it can be utilized for validation of the results from other sources.
For future studies, traffic flow can be utilized as an additional feature for constructing clusters. Also, considering other types of clustering methods can yield different results and using these methods for different bus stops can provide important outcomes which can benefit the city planners and thus the public.

Acknowledgement

This work was supported by the Bogazici University Research Fund (BAP) with the project number 11660 and the project code 16A04P2. The authors would like to thank IETT for providing bus trajectory data greatly assisting the research.

References

Chien, S. I., & Qin, Z. (2004). Optimization of bus stop locations for improving transit accessibility. Transportation planning and Technology, 27(3), 211-227.

Dueker, K. J., Kimpel, T. J., Strathman, J. G., & Callas, S. (2004). Determinants of bus dwell time. Journal of Public Transportation, 7(1), 2.

Fernandez, R. (1993). An expert system for the preliminary design and location of high-capacity bus-stops. Traffic engineering & control, 34(11).

Gong, H., Chen, C., Bialostozky, E., & Lawson, C. T. (2012). A GPS/GIS method for travel mode detection in New York City. Computers, Environment and Urban Systems, 36(2), 131-139.

Koshy, R. Z., & Arasan, V. T. (2005). Influence of bus stops on flow characteristics of mixed traffic. Journal of transportation engineering, 131(8), 640-643.

Li, M. T., Zhao, F., Chow, L. F., Zhang, H., & Li, S. C. (2006). Simulation model for estimating bus dwell time by simultaneously considering numbers of disembarking and boarding passengers. Transportation Research Record: Journal of the Transportation Research Board, (1971), 59-65.

Lin, W. H., & Zeng, J. (1999). Experimental study of real-time bus arrival time prediction with GPS data. Transportation Research Record: Journal of the Transportation Research Board, (1666), 101-109.

Mazloumi, E., Currie, G., & Rose, G. (2009). Using GPS data to gain insight into public transport travel time variability. Journal of Transportation Engineering, 136(7), 623-631.

Milkovits, M. (2008). Modeling the factors affecting bus stop dwell time: use of automatic passenger counting, automatic fare counting, and automatic vehicle location data. Transportation Research Record: Journal of the Transportation Research Board, (2072), 125-130.

Oliver, M., Badland, H., Mavoa, S., Duncan, M. J., & Duncan, S. (2010). Combining GPS, GIS, and accelerometry: methodological issues in the assessment of location and intensity of travel behaviors. Journal of Physical Activity and Health, 7(1), 102-108.

Padmanaban, R. P. S., Vanajakshi, L., & Subramanian, S. C. (2009, June). Estimation of bus travel time incorporating dwell time for APTS applications. In Intelligent Vehicles Symposium, 2009 IEEE (pp. 955-959). IEEE.

Pinelli, F., Calabrese, F., & Bouillet, E. P. (2013, October). Robust bus-stop identification and denoising methodology. In Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on (pp. 2298-2303). IEEE.

Saka, A. A. (2001). Model for Determining Optimum Bus-Stop Spacing in Urban Areas. Journal of Transportation Engineering, 127(3), 195-199.

Stoll, N. B., Glick, T., & Figliozzi, M. A. (2016). Using High-Resolution Bus GPS Data to Visualize and Identify Congestion Hot Spots in Urban Arterials. Transportation Research Record: Journal of the Transportation Research Board, (2539), 20-29.

Stopher, P. R., Jiang, Q., & FitzGerald, C. (2005). Processing GPS data from travel surveys. 2nd international colloqium on the behavioural foundations of integrated land-use and transportation models: frameworks, models and applications, Toronto.

Turnquist, M. A. (1982). Strategies for improving bus transit service reliability (No. DOT/RSPA/DPB50-81/27Final Rpt.).