EVALUATION OF THE DELAYS DUE TO ROAD CLOSURES ON ARTERIALS

Ilgın GÖKAŞAR1, Kaan AYTEKİN2

ABSTRACT

Road closures happen often in traffic due to incidents such as roadwork, maintenance and traffic accidents. These closures create bottlenecks in the traffic flow and cause considerable delay experienced by the road users. If the impact and the characteristics of a closure could be estimated beforehand, planned closures can be scheduled in a more effective way. In this study, the impact of road closures in arterials in İstanbul, obtained through one-month observation of IBB Traffic Intensity Map, is investigated. These data consist of the time and duration of the closure, the length and average speed of the upstream and downstream queue created by the closure, number of lanes in the initial conditions, and number and position of closed lanes. In the analysis, it is observed that data gathered from off peak traffic consist of two clusters. The clustering is performed using hierarchical clustering with Ward's method. The number of clusters is determined considering the silhouette scores. It is observed that data in the first cluster shows delay times with low variance and the second cluster shows a high variance; thus, unpredictability in delay times. The difference between the clusters is further investigated. It is shown that data inside the first cluster had closure times no more than 30 minutes while the second cluster had much longer closure durations. The outcome of this study showed that off peak road closures, which are taking longer than 30 minutes, result in unpredictable delay times.

Key Words: Hierarchical Clustering, Road Closure, Traffic Congestion, Traffic Delay

INTRODUCTION

Transportation network of a metropolitan city, such as Istanbul, experiences road and lane closures on a regular basis. These closures may happen due to driver errors, such as accidents or breakdowns or planned road maintenance projects. Regardless of the cause, road and lane closures affect the traffic in a disruptive manner, causing traffic bottlenecks, delays, excess fuel consumptions and increased emissions. Many studies are available which evaluate the impact of such closures using various techniques.

Decrease in the capacity of a road during lane closures for various combinations (closing 2 out of 3 lanes, 1 out of 2 lanes, 3 out of 5 lanes, etc.) is investigated in different sites of Texas by Dudek and Richards (1982) [1]. Their findings were used in the Highway Capacity Manual 1994 [2]. Texas Transportation Institute also developed a DOS (Disk Operating System)-based program called Queue and User Cost Evaluation of Work Zones (QUEWZ) that uses the equations in Highway Capacity Manual [3]. Maze et. al. (2000) did a statistical capacity analysis of work zone lane closures in Ohio [4]. Kim (2001) developed a regression model for work zone capacity estimation with explaining variables such as number of closed lanes, percentage of heavy vehicles, grade and work intensity [5]. Regression results found to be better at predicting the actual work zone capacity than the applications of Highway Capacity Manual. Li et. al. [6] conducted a data analysis to determine the work zone lane capacities along multilane corridors. 5 days of 5-minute interval data from 7 work zones is used in the study. The results showed that shifted log-normal distribution fits well to the capacity of work zones. Du et. al. [7] trained a single layer artificial neural network which takes normal speed of road segments, work zone length, work zone duration, work zone starting time and open to closed lane ratio and finds the estimated speed. The total delay is calculated with the normal and estimated work zone speeds.

1Ilgın GÖKAŞAR, Boğaziçi University Engineering Faculty, Civil Engineering Department, Istanbul, Turkey, ilgin.gokasar@boun.edu.tr
2Kaan AYTEKİN, Boğaziçi University Engineering Faculty, Civil Engineering Department, Istanbul, Turkey, kaan.aytekin@boun.edu.tr

It is seen that impacts change with respect to many aspects of the road. Road type, initial number of lanes, number of closed lanes, time of the closure are some of the aspects used in the literature. The impact also changes from city to city, depending on the characteristics of the driver population in the city. The traffic comprises of individual driving agents and their interactions in a micro level cause macro level properties. Thus, the small characteristic differences can create different results in each city. This micro level changes require an analysis to be conducted on the İstanbul traffic rather than adopting findings from other cities.

In this study, the impact of road closures in arterials in İstanbul, obtained through one-month observation of IBB Traffic Intensity Map, is investigated.

METHODOLOGY

Various data analysis tools and statistical tests are utilized for the construction and justification of clusters. The details of each step are explained in the following sub-chapters.

Data Collection and Analysis

Data collection is done manually due to the lack of API for the IBB Traffic Intensity Map to detect and record lane closures. A group of volunteer students are trained to collect relevant information in case of a lane closure. Students were trained to record the date, time and duration of the closure, number of lanes in the initial conditions, and number and position of closed lanes and the length and average speed of the upstream and downstream queues created by the lane closure with 15-minute intervals. Collected data is a time series data of resulting queues and also contains the physical information of the road (e.g. Number of lanes in the initial conditions).

The collected data is further processed. Entries with errors are eliminated and entries are grouped according to rush hour and standard (non-rush hour) entries (an entry is defined as in rush hour if the incident starting time is between 6-9 AM or 4-7 PM). Delay time is calculated for each entry as the difference between the average time spent on the length of queue and the time to pass the length of queue in the normal road conditions. Duration of the lane closure and the duration until the dissipation of the formed queue and closure ratio (number of closed lanes divided by number of lanes initially) are calculated. The rest of the analysis methodology is applied on the so-called standard time data.

Delay times of the standard times are further investigated. The histogram of delays in Figure 1 shows that the delay distribution does not resemble a normal distribution. Shapiro-Wilk normality test is applied on the delay values and a p-value of 5.442×10^{-10} is obtained, which implies the null hypothesis of normality can be strongly rejected. The same test is also applied on the closure durations and a p-value of 2.792×10^{-10} is obtained, giving the same conclusion of non-normality.

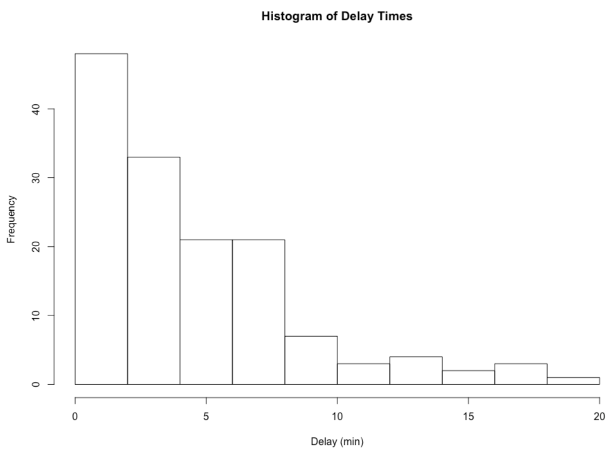

Figure 1. Histogram of Delay Times

Hierarchical Clustering and Logistic Regression

Hierarchical Clustering with Ward's Method and Euclidean distance is applied on the data. The appropriate amount of clusters is determined by looking at the average silhouette values for different cluster counts. Figure 2 shows that using 2 clusters for clustering is an adequate decision.

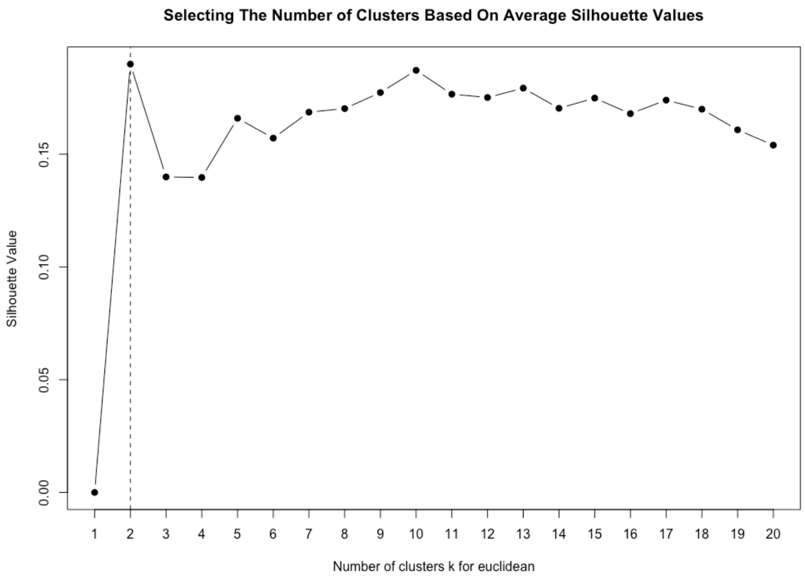

Figure 2. Silhouette Values for Each Cluster Count

After the cluster labels are obtained, the discriminating feature of clusters are investigated with logistic regression. The logistic regression is used due to its robustness against the non-normal distribution of the delays, which violates the assumption of normality of discriminant function analysis. Cluster information is used as dependent variable and all the other features of the data are used as the independent variables. As a result, delay and closure duration are found to be the only significant features with p-values of 1.57×10^{-3} and 5.73×10^{-5} respectively. The logistic regression results showed that two clusters are separated based on the delay and closure duration values. The difference in delay values can also be observed by inspection in the Figure 3. Cluster 1 shows smaller variance and smaller delay values than cluster 2.

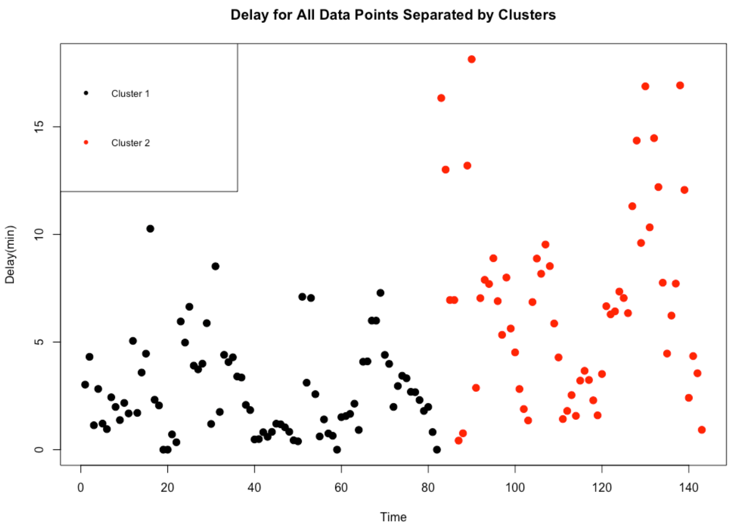

Figure 3: Delay Times for All Data Points Separated According to Clusters

Statistical Tests for Resulting Clusters

The difference in the distribution of delay and closure durations for the clusters are visualized and statistical tests are applied to see if the differences are statistically significant. Figure 4 and Figure 5 show the distribution characteristics of the delay and closure durations for both clusters, respectively. Figure 4 shows that closure durations in the first cluster rarely exceeds 30 minutes except some outliers while the cluster 2 has closure durations larger than 30 minutes. Figure 5; on the other hand, shows that delay values of cluster 2 has higher variance and a higher mean value than cluster 1.

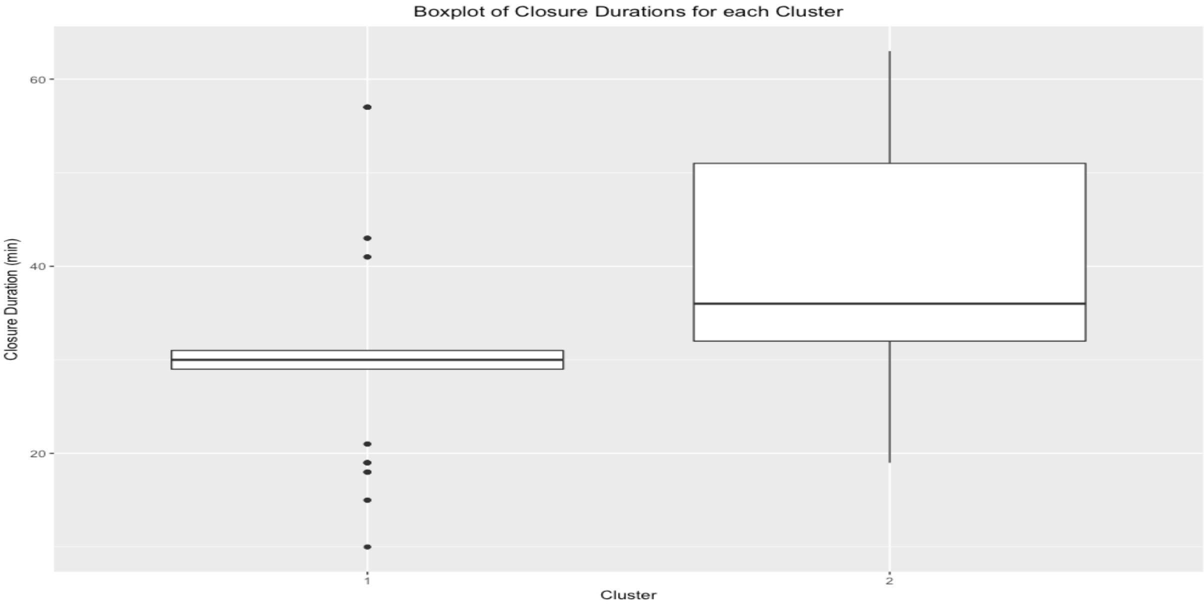

Figure 4: Boxplot of Closure Duration for Each Cluster

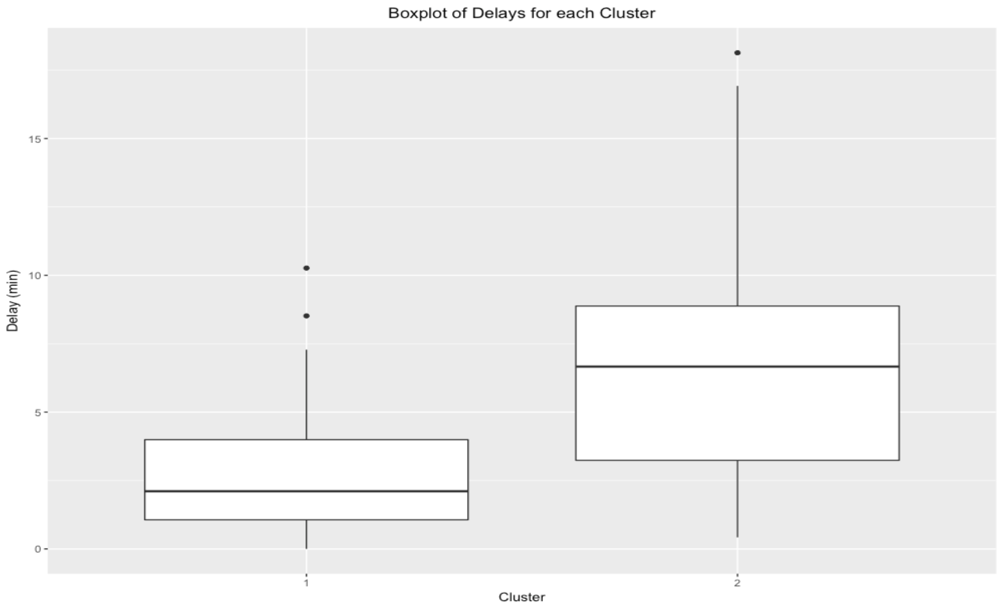

Figure 5: Boxplot of Delays for Each Cluster

After visual inspection, mean and variance of the delay and closure duration values are subjected to statistical testing. Levene's test is applied on the delay time of clusters, due to their non-normal nature, to see if their variances differ.A p-value of $2.548x10^{-6}$ is obtained, rejecting the null hypothesis of variance homogeneity. Levene's test is also applied on the closure durations and a p-value of $4.203x10^{-7}$ is obtained, again rejecting the variance homogeneity. After being aware of the variance heterogeneity, and non-normality of the data, Kruskal-Wallis test is applied on both delay and closure duration values of clusters and p-values of 6.422×10^{-10} and 2.207×10^{-8} are obtained, rejecting the null hypothesis of equal median values. These findings show that the clusters are statistically significantly different from each other in the aspects of delay and closure durations.

CONCLUSION

In this study, novel data using the IBB Traffic Intensity Map about the lane closure events was gathered. Afterwards, a data analysis on the collected data was conducted. The hierarchical clustering results showed that non-peak hour lane closures show a significantly higher mean value and a significantly higher variance in delays if the lane closing event exceeds 30 minutes. This phenomenon can be used to plan the lane closures to impact the traffic as small as possible by dividing the closure times to 30-minute intervals with enough separation time in between or charging the lane closure events in a higher rate if they exceed the 30-minute threshold. On another aspect, dividing the closure duration to intervals increases the setup time and thus setup cost. Future studies should focus on defining and calibrating lane closure cost functions depending on delay times and other related cost functions. With those findings an optimal planning model that minimizes the total cost can be developed to plan the work zone schedule with further work.

ACKNOWLEDGEMENT

This work was supported by the Bogazici University Research Fund (BAP) with **the project number 14021 and the project code 18A04P1**.

REFERENCES

[1] Dudek, C. L., Richards S. H., 1982, “Traffic capacity through urban freeway work zones in Texas” No. 869.
[2] Transportation Research Board, 1994, “Highway Capacity Manual”, Washington D.C..
[3] Krammes, Raymond A., et al, 1993, "User’s Manual for QUEWZ-92".
[4] Maze T., Schrock S. D., Kamyab A, 2000, "Capacity of freeway work zone lane closures." work 6.8 : 12.
[5] Kim, Taehyung, Lovell D. J., Paracha J., 2001, "A new methodology to estimate capacity for freeway work zones.", 80th Annual Meeting of the Transportation Research Board, Washington, DC.
[6] Li, Mingxin, Faghri A., Fan R., 2017, "Determining Work Zone Lane Capacities Along Multilane Signalized Corridors."
[7] Du, Bo, et al., 2016, "Artificial neural network model for estimating temporal and spatial freeway work zone delay using probe-vehicle data.", Transportation Research Record: Journal of the Transportation Research Board 2573: 164-171.