An automatized signal timing detection at signalized intersections with detailed vehicle trajectory data

Jasso Espadaler Clapés*[1], Emmanouil Barmpounakis[1], Nikolas Geroliminis[1]

[1] Urban Transport Systems Lab (LUTS), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

SHORT SUMMARY

Signalized intersections are a fundamental part of urban networks. Their understanding is crucial to identify congestion patterns, queues, delays, and safety issues in local and network level. In this work, we analyze multimodal vehicle trajectories and propose a methodology to extract the signal timing schedule of an intersection using the pNEUMA dataset. In addition, we combine the available information from OpenStreetMap (OSM) to map match the trajectories to the underlying network and to identify more accurately the location of traffic signals. Then the methodology to extract the signal timing schedule of an intersection consists of the following steps: i) critical movements identification, ii) computation of crossing times at the traffic signals, iii) cycle length detection and iv) phase length of each critical movement. Results show that by using the OSM data, the methodology can then be applied to any intersection in the network and provide critical information in macroscopic level.

Keywords: traffic flow theory; traffic signal timing detection; drone trajectory data.

1. INTRODUCTION

The analysis of signalized intersections in a congested urban environment is an important step towards an efficient and intelligent traffic management. However, the lack of data in such an environment has prevented the transportation researchers from analyzing intersections with real and accurate data to study their performance in terms of level of service or safety. Recently, the use of drones has allowed the collection of a massive dataset nicknamed pNEUMA. The dataset is a first-of-its-kind experiment and comprises thousands of naturalistic trajectories in central Athens, Greece (Barmpounakis and Geroliminis, 2020). This unique environment with heavy congestion and a high density of intersections challenges the existing methods of traffic analysis and needs new methods for proper analysis. Using this dataset and OSM, the aim of this paper is to propose a methodology to extract the signal timing schedule of an intersection (cycle length and phases' lengths) to allow a more detailed analysis regarding level of service, emissions, and road safety. The fact that signal timing might not always be available for use highlights the importance of having an automated method to extract it from vehicle trajectories. While the city of Athens operates with pre-timed traffic signals, we expect that this method is directly applicable for adaptive signals as well.

Signal timing detection from trajectories has not gathered a lot of attention from the traffic research community. Many intersections equipped with smart traffic signals software can already report signal timing data. However, this information is not always available to researchers or practitioners. The different methods proposed in the literature depend greatly on the input data and assume that queues are fully discharged at each cycle or oversaturated conditions. While cycle length can also be estimated in undersaturated conditions, the green time estimations would only be a lower bound of the real phases. A first example of a signal timing method is described in

(Axer and Friedrich, 2016), where the authors identify the cycle length from low frequency floating car data by checking the distribution of the times the vehicles cross the stopping line. Their main idea is to check different cycle lengths, perform a modulo operation and check if the resulting distribution resembles a uniform distribution with an Anderson-Darling test. If the cycle is correctly chosen, the distribution of crossing times should not be a uniform distribution during the whole cycle length, as vehicles are only allowed to cross during green time. Moreover, regarding the evaluation point of the crossing times the authors argue that the stopping line corresponds to the point with highest density of speeds below 3 km/h.

The stop line where to evaluate the crossing times is a vital step towards a correct signal timing detection. This can be manually set like in (Hao et al., 2012) or can be obtained from shockwave analysis (Zhou et al., 2021). It is clear that manually setting the stop lines in large urban networks with a high density of signalized intersections is not possible. On the other hand, shockwave analysis is a good tool to detect stop lines but additional information like intersection type is needed. With regards to (Hao et al., 2012), travel times before and after the intersection are used to compute cycle lengths and phases. An iterative process is used to estimate cycle breaks, effective green and red times. Regarding (Zhou et al., 2021) it represents a first approach to signal timing detection using the pNEUMA dataset. The authors used shockwave theory to analyze trajectories in a single arterial containing three signalized intersections. The cycle length detection was based on a clustering algorithm (DBscan) of the stopped vehicles, where each cluster represented a cycle. To infer the green and red times, the minimum and maximum times in each cluster were taken as a reference. While the results presented are clearly promising, the approach lacks flexibility and cannot be easily generalized to all the intersections of the network.

It can be seen that with the further development of new urban data collection techniques, new methodologies are needed to extract signal timing data in a scalable way. In the case of the pNEUMA dataset, the only information available is that traffic lights have a fixed timing plan during the recording period. Additionally, the OSMnx Python package provides an easy tool to retrieve crucial information to represent the network with nodes and edges as well as traffic signal location (Boeing, 2017). Then, it is possible to detect the intersections of the network and classify them into signalized, unsignalized or roundabouts and focus on the analysis of the first.

In this paper we propose a methodology that uses detailed and high-quality trajectory data and the available information in OSM for an efficient applicability and generalization. By retrieving the traffic signals' locations from OSM we can automatically establish stop lines to evaluate crossing times. The analysis of the crossing times is the basis of the methodology presented in the next section.

2. METHODOLOGY

In this section we provide the methodology to detect the signal timing schedule of an intersection. We will then extend the method for all intersections across the network, so to additionally identify offsets. Figure 1 shows a flowchart of the process to detect signal timing from trajectory data and OSM. The two inputs are used to detect the location of the intersections as well as the traffic signals' locations. Using a Hiden Markov Model (HMM) based map matching algorithm (Meert and Verbeke, 2018), we match the trajectories to the underlying network which is represented as a graph. This is a critical step since the map matching allows to keep track of the streets where vehicles are present, and thus it allows to check the critical movements for each direction and intersection. Once the critical movements are identified, the final step concerns the core part of this paper which described the methodology to identify cycle and phases' lengths with time-series analysis.

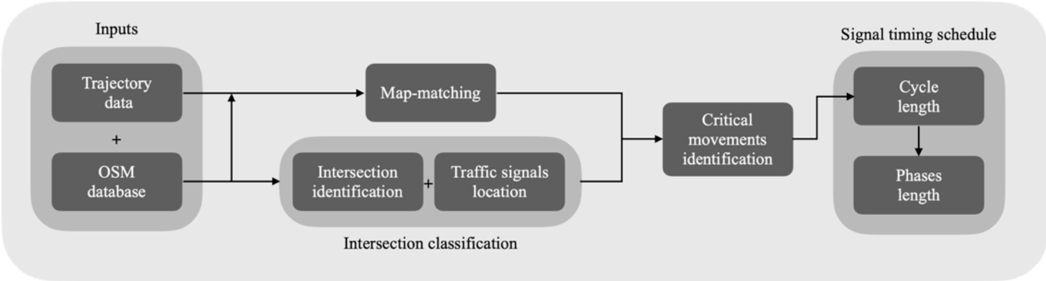

Figure 1: Methodology flowchart

An intersection is composed of two or more streets. In a simple intersection with just two roads the maximum number of movements is 12 considering that both roads are two-way, and all possible turns are allowed. Although, in a dense urban environment like the central district of Athens, intersections have limited movements allowed and numerous shapes, the methodology presented here is suitable for all kinds of signalized intersections. The added value of this methodology is that only relies on trajectory data and open Street maps, meaning that it can easily transferred to multiple locations with similar data collection.

Critical movements, stop lines and crossing times

The intersection is defined as a subgraph of the network graph. The subgraph is composed of a collection of nodes and edges, and the attributes of the nodes allow to detect the presence of traffic signals. Thanks to the map matching algorithm, each trajectory is described with the links of the network graph. Therefore, by checking the paths followed by the vehicles' trajectories we conclude which are the critical movements. However, due to noisy trajectories or illegal movements -mostly Powered-Two-Wheelers (PTWs)- the map matching algorithm can sometimes provide erroneous link assignments. Together with the fact that a big proportion of PTWs do not strictly respect traffic signals (for instance, crossing seconds before the light turns green, stopping after the stopping line etc.), it was decided to remove the PTWs from the analysis.

For each critical movement identified, the presence of traffic signals along the movement is verified. If a signal is detected a virtual loop detector perpendicular to the direction of the street is installed in its location. This allows to check the times where trajectories cross the virtual loop detector. Mathematically, we represent the set of critical movements as M and S_m as the set of vehicles driving through critical movement $m \in M$. Then, t_i^m represents the crossing time of the vehicle $i \in S_m$ in movement $m \in M$. For each movement, we obtain the vector of crossing times:

$$\overrightarrow{t^m} = [t_i^m \;\; i \in S_m], \qquad m \in M \tag{1}$$

Cycle length detection

The goal of the cycle length detection is to identify the cyclic behavior of the crossing times. In most intersections all movements have the same the cycle. However, it has been seen that in some intersections some movements have different cycle lengths. For instance, a certain movement can have a cycle of 45 seconds and other movements have a cycle of 90 seconds inside the same intersection. Therefore, c^m represents the cycle length of movement $m \in M$.

In order to obtain the cycle length of each movement the histogram of the crossing times is computed in 1-second bins. Then, a smoothing of the histogram with a moving average (20-seconds window size) is performed. This allows to detect crossing time peaks that repeat every cycle. To ease the peak detection task, the autocorrelation function of the smoothed histogram is also computed. The autocorrelation function allows an easier peak detection since all the values are always in the interval [-1, 1]. Then, using a peak detection algorithm with a minimum distance of 40 seconds between peaks (equivalent to a minimum cycle constraint) the times of the peaks are identified. Then, p_j^m represents the time of the j-th peak for movement $m \in M$, and the difference between two neighboring peaks is equal to the cycle length detected. Finally, $\overrightarrow{c^m}$ represents the vector of cycles detected in movement *m*, and the median of the cycles' distribution is assumed to be the cycle length c^m of movement *m* (rounded to the nearest multiple of 5).

$$\begin{gathered} \overrightarrow{c^m} = \left[p_{j+1}^m - p_j^m \quad 0 \leq j \leq N-1\right] \\ c^m = median(\overrightarrow{c^m}) \end{gathered} \tag{2}$$

Phase length detection

As aforementioned, the cycle length is only a value indicating the repetitiveness of the crossing times. To complete the signal timing the red and green phases need to be detected. To do so, we take the reminder of the division between the crossing times t_i^m and the cycle length c^m, also known as modulo operation (Equation (3)). Then every crossing time is transformed such that:

$$t_{i,c}^m = t_i^m \; mod \; c^m \tag{3}$$

Performing the modulo operation (expressed as mod) allows to see the real crossing times inside a cycle. By analyzing the distribution of the new crossing times $t_{i,c}^m$, the 1st and the 99th percentiles are assumed to be the starting and ending times of the effective green phase g^m. Although PTWs have been removed from this calculation, the choice of the 1st percentile reflects the impatience of some drivers to cross the traffic signal when it is about to turn green. In a similar way, the choice of the 99th percentile is also justified by those vehicles that cross the traffic signal moments after turning red. The yellow phase is assumed to be 3 seconds and finishes with the ending of the effective green time.

3. RESULTS AND DISCUSSION

The results shown in this section refer to the intersection between Alexandras Avenue and October 28th Avenue, one of the busiest intersections in central Athens. Figure 2 shows the outline of the intersection from OSM on the left side, while the right side presents the graph representation. As aforementioned, extracting the OSM network and its characteristics allows to represent intersections as graphs and identify the location of traffic signals (marked in red on the graph).

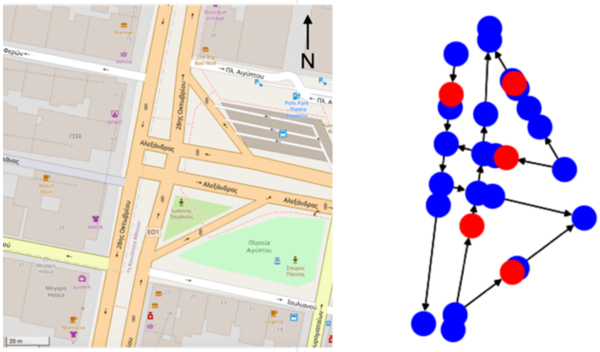

Figure 2: Intersection between Alexandras Avenue and October 28th Avenue. The layout from OSM is shown on the left while the graph representation on the right

Next, we show an example of the methodology described in section 2 for a specific movement of the intersection in Figure 2 (movement number one concerns the flow of vehicles driving from Alexandras Avenue and turn left towards October 28th Avenue). The trajectories used for the purpose of this example are from Monday 29th October 2018 from 10:00 to 10:30. Figure 3a shows the histogram of the times vehicles crossed the stop line of the traffic signal. A clear cyclic behavior can be identified from the histogram. To recognize the cycle length of this movement, Figure 3b shows the moving average of the histogram and Figure 3c) shows the autocorrelation function of the moving average. The peaks of the autocorrelation function have been added on top with orange crosses, and after analyzing them the cycle of this movement is set to 90 seconds. Finally, Figure 3d presents the histogram of the modulo times inside the cycle. The 1st and the 99th percentiles indicate the start and the end of the effective green, which has been set to 24 seconds.

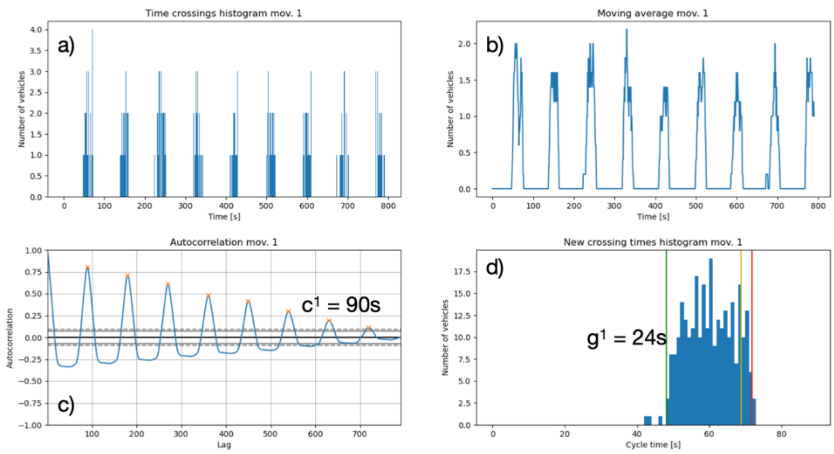

Figure 3: Example of the methodology for a particular movement. A) Histogram of time crossings, b) Moving average of the histogram, c) Autocorrelation function, d) Histogram of times inside cycle

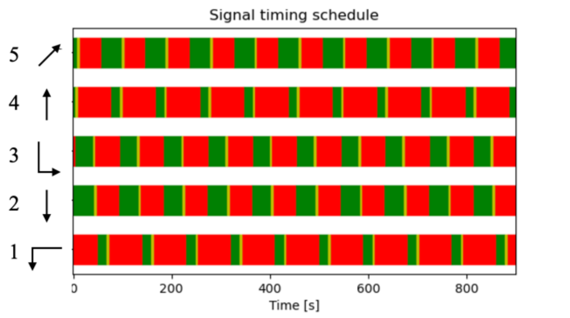

Figure 4: Signal timing schedule for the critical movements of the intersection

The methodology was repeated for every critical movement identified in the intersection, which allowed to identify all the green times and their sequence in time. This is shown in Figure 4, where the signal timing schedule with the critical movements is presented. For this intersection, the analysis of the trajectories showed a perfect cyclic behavior since all the movements had a cycle length of 90 seconds. The phases for movements 1, 2, 3, 4 and 5 were 24, 44, 39, 22, and 46 seconds respectively.

The results regarding the intersection of Figure 2 refer to a saturated intersection with no red-light violations. However, the high density of intersections in the pNEUMA dataset presents a variety of situations that differ from the intersection of the example. For example, small intersections in under-saturated conditions are difficult to analyze especially regarding the phases' times. Indeed, the results provided by the methodology are only a lower bound of the real green times. Figure 5 shows the results of the methodology when applied to the 97 signalized intersections of the study area of the pNEUMA dataset (data from 01/11/2018, 10:00-10:30). The data reveals that most intersections have 90-seconds cycles while some intersections in specific arterials have 45-seconds cycles. Moreover, some intersections have combined cycles (45 and 90 seconds). These have been plotted with two colors.

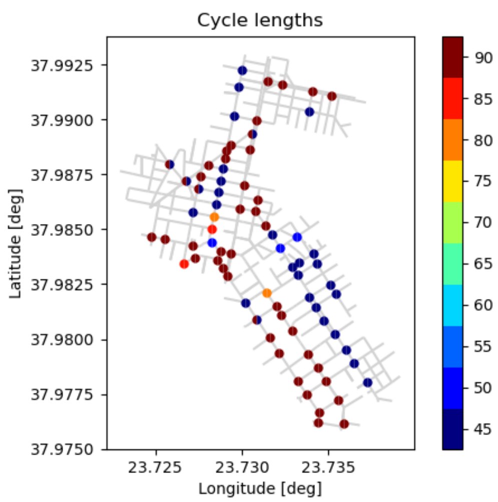

Figure 5: Cycle lengths of the intersections in central Athens

Finally, Figure 6 presents a time-space diagram for a 700-meters segment of Panepistimiou Avenue, one of the most heavily congested in Athens. The diagram also shows the green, yellow and red times of the through movements at the 7 intersections of the segment. The results show the synchronization of the traffic signals along the arterial, which allows most of the vehicles to drive through the avenue without stopping at the intersections. In addition, an increased flow during yellow phases can be observed at some intersections.

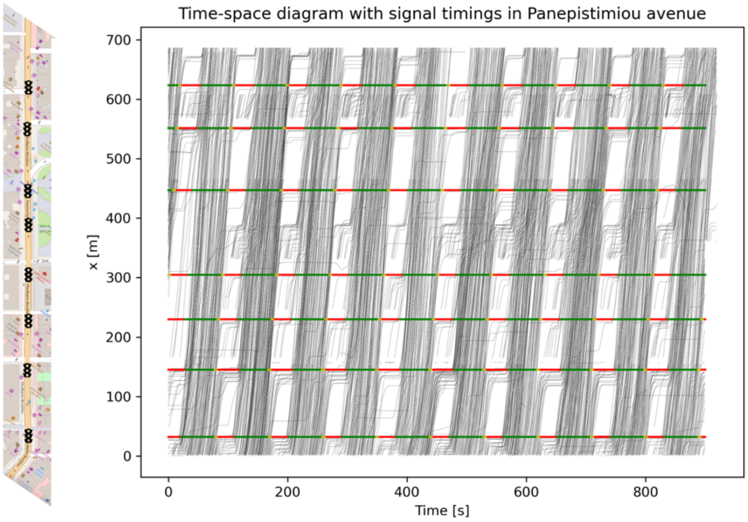

Figure 6: Time-space diagram in Panepistimiou Avenue for through movements with green and red times at the intersections

4. CONCLUSIONS

The research presented in this short paper proposes a new methodology to extract signal timing from high-quality vehicle trajectories using powerful data analysis tools as well as available information in OSM. The representation of intersections as graphs allows to detect critical movements and the location of traffic signals from OSM provides an evaluation point to check the crossing times. The analysis of the crossing times -the basis of the methodology- provides crucial information like cycle length, phase time and a full signal timing schedule.

The methodology presented here is scalable to any signalized intersection of the network. The computation of the cycle length of each movement is particularly stable since it is only a measure of the repetitiveness of the crossing times. On the other hand, the calculation of the phase times requires saturated conditions to be accurate. In the case of undersaturated conditions, the estimation of the green time is only a lower bound of the real phase time.

ACKNOWLEDGMENT

The authors would like to acknowledge the partial funding supported by the EPFL Open Science Fund and Swiss National Foundation project 200021_188590.

REFERENCES

Axer S, Friedrich B. A Methodology for Signal Timing Estimation Based on Low Frequency Floating Car Data: Analysis of Needed Sample Sizes and Influencing Factors. Transportation Research Procedia 2016;15:220–32. https://doi.org/10.1016/j.trpro.2016.06.019.

Barmpounakis E, Geroliminis N. On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment. Transportation Research Part C: Emerging Technologies 2020;111:50–71. https://doi.org/10.1016/j.trc.2019.11.023.

Boeing G. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Computers, Environment and Urban Systems 2017;65:126–39. https://doi.org/10.1016/j.compenvurbsys.2017.05.004.

Hao P, Ban X (Jeff), Bennett K, Ji Q, Sun Z. Signal Timing Estimation Using Sample Intersection Travel Times. IEEE Transactions on Intelligent Transportation Systems 2012;13. https://doi.org/10.1109/TITS.2012.2187895.

Meert W, Verbeke M. HMM with Non-Emitting States for Map Matching 2018.

Zhou Q, Mohammadi R, Zhao W, Zhang K, Zhang L, Wang Y, et al. Queue Profile Identification at Signalized Intersections with High-Resolution Data from Drones. 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2021, p. 1–6. https://doi.org/10.1109/MT-ITS49943.2021.9529337.