Simulating Indiana's S 7th Street Six-Way Intersection

Morgan Buterbaugh, Alan Davis, Akinbusola Olushola, Sinti Sasai, Toby Trotta

Department of Mathematical and Computer Sciences

Indiana University of Pennsylvania

Abstract

To enhance transportation flow at street intersections, it is imperative to establish proper coordination among the signals and traffic lights that manage the flow of traffic in multiple directions. We propose a study that presents a methodology to assess impractical traffic patterns at a six-way intersection in Indiana, causing unpredictable wait times and delays in arrival times for local destinations and to determine optimal signal timings to minimize traffic congestion. In representing this system as close to reality, we will compile data to evaluate how the system currently is. This will include collecting such data as vehicle and pedestrian count at this intersection, at different time cycles, and data shown by google maps. Information relevant to the arrival and vehicles in that area will be collected and proceeded by finding the best probability distribution of the data according to the type of vehicle and route where it entered. For the simulation model, our goal is to find the best random distribution associated with the time between the arrival of the vehicles in the rush hour segment of analysis using the goodness of fit test. We aim to optimize the model and determine the necessity of all streets for travel, identifying which turning lanes are necessary, and evaluating the effectiveness of different traffic control strategies and infrastructure improvements, that will help in achieving our objective.

1 Introduction

Traffic congestion at intersections is a common problem that can lead to long wait times, delays, reduced safety for both drivers and pedestrians, reduced productivity, increased air pollution, and increased fuel consumption. One of the major causes of traffic congestion is the mismanagement of intersections, leading to delays and unpredictable wait times for drivers and pedestrians. Therefore, proper coordination of signals and traffic lights is crucial to enhance transportation flow at street intersections. To address this issue, researchers have proposed a variety of approaches to optimize traffic flow at intersections. These approaches have utilized simulation models, real-time traffic data, and machine learning techniques to evaluate different traffic control strategies and improve signal timings.

One such approach is the use of simulation models, which allow researchers to analyze the performance of different signal timings and traffic control strategies in a virtual environment. Chen et al. (2021) [1] proposed a multi-objective optimization model for signal timings at a complex urban intersection in China based on an agent-based simulation platform. The model was able to evaluate different traffic control strategies and optimize signal timings based on multiple objectives, such as minimizing delay and maximizing throughput. Similarly, Yu et al. (2019) [3] used a micro-simulation model to optimize signal timings at an intersection in Beijing, leading to significant improvements in intersection performance.

Another approach is the use of real-time traffic data to adjust signal timings dynamically. Zhang et al. (2020) [4] developed a real-time signal optimization system that used data from connected vehicles to adjust signal timings at an intersection in Beijing. The system was able to reduce delays and improve traffic flow by adjusting signal timings in response to real-time traffic conditions. Similarly, Liu et al. (2020) [2] proposed a deep learning-based approach for signal timing optimization at an intersection in China, which used a convolutional neural network to predict traffic flow and optimize signal timings based on predicted traffic conditions.

These approaches have been successful in improving traffic flow and reducing congestion at complex intersections. The proposed study aims to evaluate traffic patterns at a six-way intersection in Indiana, PA, and determine optimal signal timings to minimize traffic congestion. By compiling data and evaluating the effectiveness of different traffic control strategies, the study will contribute to the optimization of intersection signal timings, leading to reduced delays and improved transportation flow. The study builds on previous research that has successfully used simulation models, real-time traffic data, and simulation techniques to optimize traffic flow and reduce congestion at complex intersections.

Figure 1: The six-way intersection located in Indiana, PA used for this study.

2 Data Collection

The data collection for this study consisted of observing the traffic and stoplight patterns of the six-way intersection during rush hour traffic. Data was collected on two different occasions between 3:00 PM and 5:00 PM during a normal work week.

The data we were interested in and collected were the green light times, red light times, and interarrival times. For this study, we did not focus on yellow light times and instead considered the time that the light was yellow to be green light time. To calculate how much time had passed, we used an online stopwatch.

To collect the timing of the lights, we used a stopwatch to keep track of the time and splits. For the green light times, we started the stopwatch when the light would initially turn green and then hit the 'lap' button when the light would immediately turn red. This would then tell us how long each green light lasted. Similarly, with the red lights, when the 'lap' button was hit this would measure the amount of time when the light would initially turn red. We would then hit the 'lap' button again when the light would immediately turn green. This process would then continue for a total of 31 minutes.

To collect the interarrival times of the traffic, we again used a stopwatch to keep track of the time and the splits. For each street, we chose a reference point that would represent where we would start to include the traffic in our study. Once the first car would pass the reference point we would start the stopwatch. Then, after each proceeding car would pass the reference point we would hit the 'lap' button. This process would then continue for at least 30 minutes and would tell us how long it took for each car to arrive at the corresponding light. The goal for the 30-minute collection times was to hopefully acquire a decent sample size but that was not always possible considering the popularity of the street while remaining in our proposed time limitations, 3:00 PM to 5:00 PM.

3 Summary Statistics

After all of the data was collected, the raw data was imported into an Excel spreadsheet. We then cleaned the data by doing a few things. First, we removed any errors we made by hitting the 'lap' button on accident; doing this did not mess up any of the data points. Second, we converted the time from minutes into seconds.

For the data associated with the timing of the lights, we had to label each split accordingly with either 'green' or 'red', and then we grouped the red data points and green data points together. Once we made all of these changes, we created a new Excel workbook that contained only the clean data. This workbook contained the number of seconds between each arrival, the number of seconds the traffic light was green, and the number of seconds the traffic light was red for each of the six streets. This is the data set that was used for the simulation with our ARENA model.

After the data was cleaned and correctly categorized, the summary statistic analyses were conducted using RStudio.

3.1 Interarrival Times

Let the street name be denoted by the random variables assigned in Table 3. The following summary statistics were calculated using R:

Street	Sample Size	Min.	Max.	Mean	Std. Dev.	Skewness	Kurtosis
SI	30	4	544	126.13	125.17	1.70	2.79
SO	120	2	98	15.47	15.92	2.19	6.25
WI	194	1	53	9.63	10.14	1.87	3.36
WO	65	0	186	29	34.36	2.38	7.05
LE	31	5	397	104.55	96.30	1.49	1.91
LW	22	5	254	92.18	72.04	0.69	-0.79

Table 1: Interarrival time summary statistics.

We see from the summary statistics all the data is right-skewed, which was predicted, especially for the more popular routes (South 7th Street Outgoing and Wayne Avenue Incoming). From the minimum and maximum interarrival times, sample size, and smaller values for skewness, we can deduce that South 7th Street Incoming, Locust East, and Locust West were the *least* popular streets for travel.

3.2 Green Light Times

Let the street name be denoted by the random variables assigned in Table 4. The following summary statistics were calculated using R:

Street	Sample Size	Min.	Max.	Mean	Std. Dev.	Skewness	Kurtosis
SI	15	12	24	16.60	4.40	0.54	-1.44
SO	25	12	61	36.56	14.05	-0.32	-0.88
WI	20	31	86	38.85	13.36	2.37	5.19
WO	17	9	20	16.47	2.24	-1.94	4.70
LE	15	9	16	12.40	2.47	-0.06	-1.50
LW	19	10	26	19.26	4.16	-0.62	-0.59

Table 2: Green light time summary statistics.

From the summary statistics, we can not really make any assumptions about the green light data. Both South 7th St. Incoming and Wayne Ave. Incoming are right-skewed. Both Wayne Ave. Outgoing and Locust St. West are left-skewed. South 7th St. Outcoing is normally distributed and Locust St. East is uniformly distributed. We can see from the minimum and maximum interarrival times, sample size, and smaller values for skewness that South 7th St. and Wayne Ave. Incoming are more popular routes, and Locust St. East, Locust St. West, and Wayen Ave. Outgoing are the least popular routes.

3.3 Fitted Distributions

A probability distribution is a function that is able to describe the probability of an event happening. For this project, we fitted several distributions to the collected data. As previously mentioned, we collected the interarrival times for the cars, red light (waiting process), and green light (moving process) times for each of the six streets. Figure 8 plots a histogram of the collected data, as well as a fitted probability distribution to be used in the ARENA simulation model. ARENA will consider the probability distributions as a foundation for how quickly cars will arrive at their respective lights.

Street	Random Variable	Probability Distribution
S 7th St., Incoming	SI	$SI \sim 4 + \text{Exp}(122)$
S 7th St., Outgoing	SO	$SO \sim 1.5 + \text{Logn}(16.3, 33)$
Wayne Ave., Incoming	WI	$WI \sim 0.5 + \text{Logn}(9.62, 14.9)$
Wayne Ave., Outgoing	WO	$WO \sim -0.001 + \text{Logn}(55.9, 224)$
Locust St., East	LE	$LE \sim 5 + \text{Exp}(99.5)$
Locust St., West	LW	$LW \sim 5 + \text{Exp}(87.2)$

Table 3: Vehicle Interarrival Time Fitted Probability Distributions

Using the ARENA Input Analyzer with the cleaned data, we were able to fit the best-fit probability distributions and the histograms with the superimposed probability distribution from Table 3 can be found in Figure 8. Again, these probability distributions were the foundation in constructing the ARENA Simulation Model.

Figure 15 plots a histogram of the collected data for the green light times, as well as a fitted probability distribution to be used in the ARENA simulation model. ARENA will consider the probability distributions as a foundation for how long cars have to drive at their respective lights.

Street	Random Variable	Probability Distribution
S 7th St., Incoming	SI_G	$SI_G \sim 11.5 + \text{Exp}(5.1)$
S 7th St., Outgoing	SO_G	$SO_G \sim \text{Norm}(36.6, 13.8)$
Wayne Ave., Incoming	WI_G	$WI_G \sim 30.5 + \text{Logn}(7.69, 13.8)$
Wayne Ave., Outgoing	WO_G	$WO_G \sim \text{Tria}(8.5, 17.5, 20.5)$
Locust St., East	LE_G	$LE_G \sim \text{Unif}(8.5, 16.5)$
Locust St., West	LW_G	$LW_G \sim \text{Tria}(9.5, 23, 26.5)$

Table 4: Green Light Time Fitted Probability Distributions

Using the ARENA Input Analyzer with the cleaned data, we were able to fit the best-fit probability distributions and the histograms with the superimposed probability distribution from Table 4 can be found in Figure 15. Again, these probability distributions were the foundation in constructing the ARENA Simulation Model.

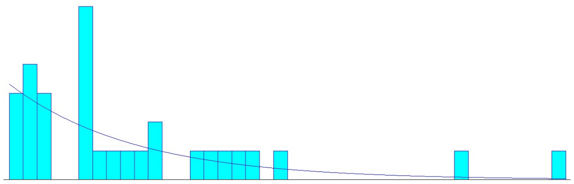

Figure 2: S 7th St. Incoming Interarrival Times Distribution

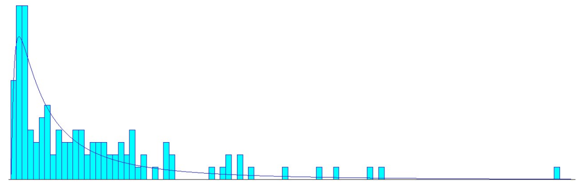

Figure 3: S 7th St. Outgoing Interarrival Times Distribution

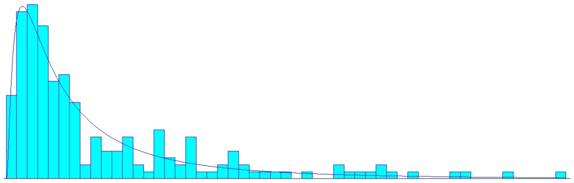

Figure 4: Wayne Ave. Incoming Interarrival Times Distribution

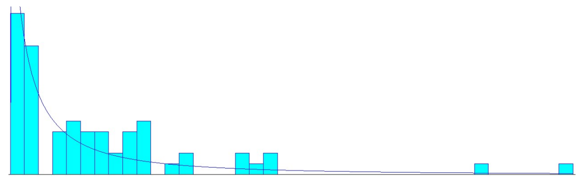

Figure 5: Wayne Ave. Outgoing Interarrival Times Distribution

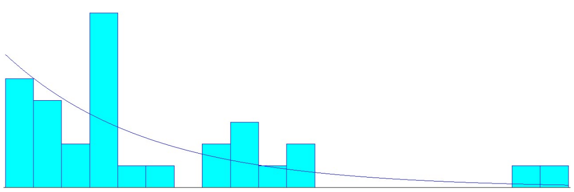

Figure 6: Locust St. (East) Incoming Interarrival Times Distribution

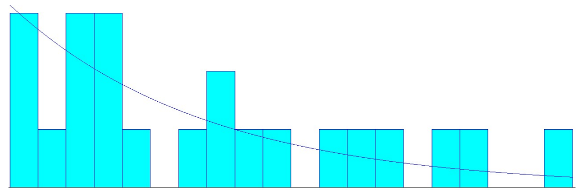

Figure 7: Locust St. (West) Outgoing Interarrival Times Distribution

Figure 8: Vehicle Interarrival Time Histograms with Superimposed Probability Distributions from Table 3.

4 ARENA Simulation Model

(Alan)

5 Simulation Summary

Due to the stochastic nature of microscopic traffic simulation (where different types of vehicles are released onto the network according to specified random distributions), we ran the model with 5 replications with time length 4 hours and confident interval of 0.95 in the Arena Software to obtain a reasonable estimate statistic summary of the results.

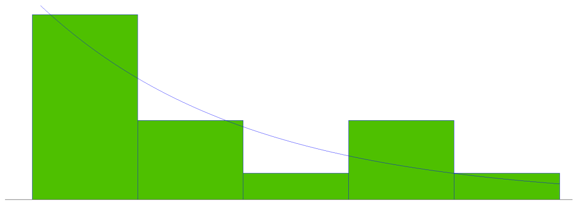

Figure 9: S 7th St. Incoming Green Light Time Distribution

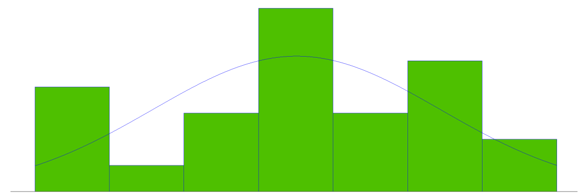

Figure 10: S 7th St. Outgoing Green Light Time Distribution

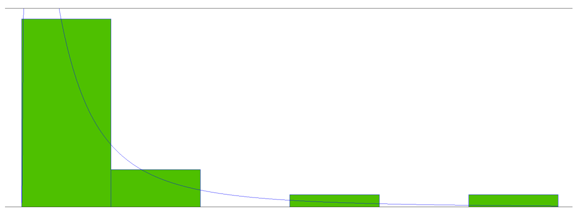

Figure 11: Wayne Ave. Incoming Green Light Time Distribution

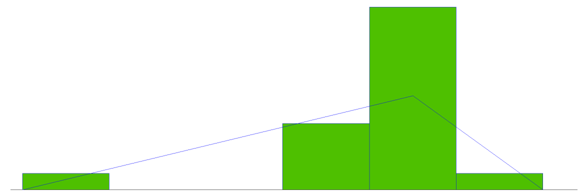

Figure 12: Wayne Ave. Outgoing Green Light Time Distribution

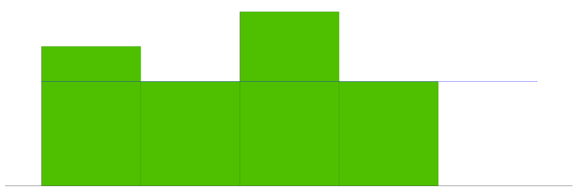

Figure 13: Locust St. (East) Incoming Green Light Time Distribution

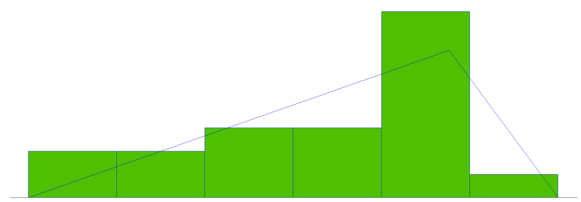

Figure 14: Locust St. (West) Outgoing Green Light Time Distribution

Figure 15: Green Light Time Histograms with Superimposed Probability Distributions from Table 4.

Tables below summarizes our traffic simulation results.

Name	Average Waiting Time	Average No. Of Observation
Light 1 Wayne	80.56	70.4
Light 2 for Locust West	41.70	141
Light 3 for South 7th North	58.98	422.2
Light 4 for South 7th South	83.63	110.2
Light 5 for Jimmy Wayne	28.12	1721.6
Light 6 Locust East	41.30	162
Light 7 for Right Lane South 7th North	61.00	444.2

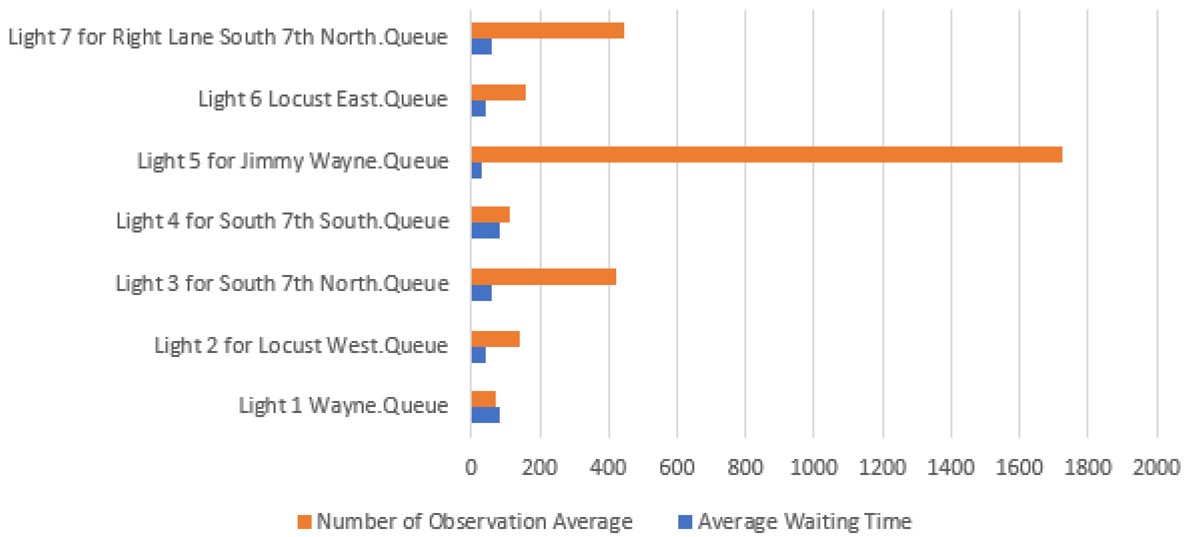

Figure 16: Traffic Simulation Results

References

[1] Jiaxiang Chen, Shasha Liu, Wei Liu, and Hai Huang. A multi-objective optimization model for signal timings at a complex urban intersection based on an agent-based simulation platform. *Transportation Research Part C: Emerging Technologies*, 123:102938, 2021.

[2] Yitong Liu, Yangsong Ye, Xiaopeng Li, and Jie Liu. Deep learning-based approach for signal timing optimization at an urban intersection. *IEEE Transactions on Intelligent Transportation Systems*, 22(6):3507–3518, 2020.

[3] Heng Yu, Shengbo Zhang, and Yong He. Optimization of signal timing at an urban intersection using a microsimulation model. *Journal of Advanced Transportation*, 2019:1–11, 2019.

[4] Qi Zhang, Zhiheng Li, Yi Li, and Xinmiao Yang. A real-time signal optimization system for connected vehicle environment based on deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, 22(6):3467–3479, 2020.