# THE FLORIDA DEPARTMENT OF TRANSPORTATION

# "Dynamic Intersection Learning Machine Optimization Real-time Engine"

## Final Report

to

THE FLORIDA DEPARTMENT OF TRANSPORTATION
TRAFFIC ENGINEERING AND OPERATIONS OFFICE

TASK 2 (Identify corridors of interest and acquire data and simulation models)

FDOT Contract BDV31-562


**Sanjay Ranka and Anand Rangarajan**
Computer and Information Science and Engineering
University of Florida

# TABLE OF CONTENTS

List of Figures

List of Tables

## 1. Introduction

Developing traffic signal timings have evolved over the years with advanced software and mathematical models [1] along with a variety of advanced traffic signal controllers that are currently available in the market. Conventional methods included using critical movement analysis to favor the high traffic approaches and synchronize adjacent intersection based on offsets and cycle lengths. Since then signal timing design and traffic signal systems have matured exponentially to provide the traffic engineer with a latitude of options (e.g. volume density function, traffic responsive plans etc.) to design advanced signal timing parameters with options to develop actuated coordinated movements to maximize operational efficiencies. Although there are multiple software available to optimize signal timing, it is time consuming and labor (as well as data) intensive. Agencies lack the staff and resources to keep up with ever change trends in traffic and there is a need to automate the process which is outcome-based signal timing approach that fits the objective needs of the agency.

Advanced signal controllers can plan for certain events or conditions in a variety of way, for example traffic responsive systems (TRS) that are able to coordinate systems and choose coordination plans that fit certain predetermined traffic conditions. Further there are several adaptive systems that have emerged in the past decade that uses detection data and algorithms to adjust signal timing parameters for existing conditions. The difference between TRS and adaptive systems is that adaptive systems can vary timings and not have prefixed plans as in TRS or other traditional signal systems based on traffic patterns. This also overcomes the coordination loss due to controller transition due to abrupt plan changes. However, it is critical to understand that there is no system that is “one size fits all” since traffic demand may vary significantly and any system including adaptive systems have prefixed local and global parameters that needs timely review at regular interval. In addition, current systems do not inherently understand the needs of every agency since each agency has varying objectives and priorities based on times of the day and under different traffic conditions. The commercially available systems have proprietary algorithms which may not interface well with open source-based systems that are deployed statewide. There is a clear and pressing need to have a state of the art traffic policy system which can be easily updated over time while being consistently responsive to traveler and pedestrian needs.

Based on the above, traditional signal systems can benefit from the latest advances in machine learning. The purpose of this research effort is not to develop a system that will assume the responsibilities of signal controller but rather explore the opportunity to use the latest machine learning methodology in making advances with existing systems. This will help deployments of signal policy plans and justify investments in new information processing hardware. This research seeks to investigate the application of machine language in signal timing optimization, explore potential challenges and propose innovative solutions towards the vision of automating signal timing adjustments. If machine learning can be successfully integrated with existing systems, it would potentially open an array of future applications not only in operational aspects but also in safety, reliability and planning (for resource allocation). With automated vehicles technology around the corner, the integration of machine learning with signal systems would provide additional capabilities especially in the transition phase where there will be a mix of AV, CV and manual vehicles creating a complex heterogeneous environment. The primary scope of this effort is to evaluate the application of machine learning in the field of traffic signal systems. This effort is novel and will be performed in a controlled environment using Hardware-in-the-loop systems (HILS) which will use a combination of simulation software and signal controller hardware. HILS is proposed since it provides the flexibility to explore and quantify system performance without any effect on the real-world traffic conditions.

This report presents the supporting tasks and deliverables for Tasks 2 of the project. The deliverables defined for this task are as follows:

*Task 2: Identify corridors of interest and acquire data and simulation models*

*Data and simulation models are required to be stored in a database for developing and testing the adapters, running simulations and understanding the type and frequency of missing information. This task will ensure that the following is available in an accessible form:*

- *Existing models developed for the Central Florida region*
- *Roadway geometry for the corridors of interest*
- *Demand profiles (volumes, turning movements, etc.) for different “base conditions” (time of day periods, days of the week, and seasons/special events)*
- *Signal timing parameters (green times, phasing, offset times, etc.) for the same “base conditions” as before*
- *Work with other researchers and practitioners to retrieve existing models for use in the rest of this project.*

*<u>Deliverable:</u> Upon completion of Task 2, the University will submit to District 5 the following:*

- *a written Technical Memorandum summarizing the findings and*
- *a set of models and other resources acquired for this task*

## 2. Existing models developed for the Central Florida region

We received VISSIM [2] based models for a city in Canada. These models were useful in understanding the type of models that can be utilized by simulators. In the Task 1 report, we compared VISSIM and SUMO (Simulation of Urban Mobility) and pointed out the advantages of the latter (open source and a wider non-proprietary ecosystem of user interface add ons). Other simulators exist, notably, AIMSUM, TSIS-CORSIM, MATSim, TRANSIMS but SUMO was picked as the simulator of choice due to its open source nature and rich ecosystem. SUMO is also a general purpose simulator which is not the case for some of the other simulators listed above. Since the city of Canada models (mentioned above) were VISSIM-based, once the decision to use SUMO was made, this model served mainly to set similar goals for developing a similar SUMO-based model in Seminole county (Florida) which is the primary goal of this work. In summary, the various decisions taken with respect to simulation software, machine learning stacks, traffic policy etc. culminated in the choice of SUMO and Python-based machine learning with an emphasis on multi-agent learning with concrete application to 329 signalized intersections in Seminole county, FL This integration will be further elaborated in the next set of tasks.

## 3. Roadway Geometry and Corridors of Interest

The real data for the project is drawn from ATSPM (Automated Traffic Signal Performance Measures) and HERE.com. These data sources are ideally suited to signal policy learning since both data sources contain information regarding vehicle arrivals on green/red signals at intersections (ATSPM) and average velocity per corridor (HERE.com) etc. Based on this information, machine learning algorithms (integrated with simpler optimization strategies) can design signal policies which aim to improve measures of effectiveness such as traffic throughput, average delay and so on. After obtaining new policies, SUMO can be utilized to simulate new traffic patterns for the set of Seminole county intersections. Subsequent policy updates are then driven by simulated data since real world experiments of this nature are difficult (if not impossible) to carry out. The basic thrust of these machine learning approaches were earlier described in the Task 1 report and are updated here using real world data (ATSPM and HERE.com).

The ATSPM dataset contains rich information about each intersection metadata:

- "Approaches" Sheet describes the various approaches per signal.
- "Signals" sheet gives physical coordinates of the signal.
- "Detectors" sheet describes individual detectors and associated signals.

| Legend | |
|---|---|
| x | Enumeration CAPTURED - not present in all controllers' outputs. |
| x | Enumeration CAPTURED - present in all controllers' outputs. |
| x | Enumeration not CAPTURED. |
| **Abbreviation** | **Description** |
| M | Mandatory (Operation and command shall be implemented) |
| O | Optional (Operation and command may be implemented) |
| D | Deprecated |

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | ApproachID | SignalID | DirectionTypeID | Description | MPH | ProtectedPhaseNumber | IsProtectedPhaseOverlap | PermissivePhaseNumber |
| 2 | 8438 | 1005 | 1 | Northbound | 45 | 2 | 0 | NULL |
| 3 | 8439 | 1005 | 2 | Southbound | 45 | 6 | 0 | NULL |
| 4 | 8440 | 1015 | 1 | Northbound | NULL | 2 | 0 | NULL |
| 5 | 8441 | 1015 | 2 | Southbound | NULL | 6 | 0 | NULL |
| 6 | 8442 | 1020 | 1 | Northbound | NULL | 2 | 0 | NULL |
| 7 | 8443 | 1020 | 2 | Southbound | NULL | 6 | 0 | NULL |

| ApproachRouteId | ApproachOrder | RouteDetailID | ApproachID |
|---|---|---|---|
| 2 | 1 | 3 | 8800 |
| 2 | 2 | 4 | 8802 |
| 3 | 5 | 18 | 8712 |
| 3 | 3 | 20 | 8709 |
| 3 | 4 | 22 | 8710 |
| 3 | 2 | 24 | 8706 |
| 3 | 6 | 26 | 8714 |

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ID | DetectorID | DetChannel | DistanceFromStopBar | MinSpeedFilter | DateAdded | DateDisabled | LaneNumber | MovementTypeID | LaneTypeID | DecisionPoint | MovementDelay | ApproachID | DetectionHardwareID |
| 2 | 12752 | 100504 | 4 | 330 | NULL | 05:02.0 | NULL | 1 | 1 | 1 | 0 | 0 | 8438 | 3 |
| 3 | 12753 | 100505 | 5 | 330 | NULL | 00:00.0 | NULL | 1 | 1 | 1 | 0 | 0 | 8438 | 3 |
| 4 | 12754 | 100510 | 10 | 330 | NULL | 00:00.0 | NULL | 1 | 1 | 1 | 0 | 0 | 8439 | 0 |

Figure 1 ATSPM Intersection Related information

The data intensive nature of traffic intersections (and elsewhere) calls for a comprehensive investigation into data mining, machine learning and integration with traditional methods (actuated control, optimization). In the Task 1 report, the advantages and disadvantages of modern data and information processing methods relative to traditional methods were highlighted. Here, these are summarized. Changes to signal policy plans directly affect traffic behavior which in turn can be analyzed leading to more signal policy modifications. This may seem circular but is ideally suited for machine learning (and reinforcement learning) since policy updates can be tuned to real word changes in an iterative modification loop. Traditional optimization methods are not as well suited for this purpose and have focused more on single intersections rather than on the entire network.

Based on availability of data and input from sponsor, we decided to focus on 329 Signalized Intersections from Seminole County, D5, Florida (Figure 2). The initial data ranges for dates between 1st May 2016 to 16th September 2016 for ATSPM and Here.com. The total storage space occupied is ~ 400 GB and consists of several comma-separated value files. We also received this data for later part of 2018.


Figure 2: Map of Seminole county with all the intersections highlighted

## 4.0 Demand profiles and Signal Timing Parameters

As discussed earlier, we decided to focus on ATSPM and HERE.com datasets for deriving the demand profiles and signal timing parameters for the corridors of interest. These datasets are described in detail below for achieving these objectives.

### 4.1 ATSPM Data

Automated Traffic Signal Performance Measures (ATSPM) [3] data is a stream of data obtained by modern traffic intersection signal controllers. Induction loop detectors attached to the intersection collect data at 10 Hz, indicating whether a vehicle passed over it or not. Signal behavior is also captured. This data allows traffic engineers to analyze the performance of traffic intersections and improve safety and efficiency while cutting costs and congestion. Below we describe the data in great detail. It is worth pointing out at this juncture that this dataset can be utilized for more than signal policy optimization. For example, often detector lane mappings are not available and must be reverse engineered from the raw ATSPM dataset.

**Raw Data Files**: 22 such comma-separated value (.csv) files are available. Each is between 10-17 GB each and contains the data recorded at decisecond frequency. Each file contains about a week of raw data. The data consists of 4 columns:

- SignalID
- Time of recording
- EventCode: What event at the signal was captured
- EventParam: What was the value of the event or attribute at that timestamp

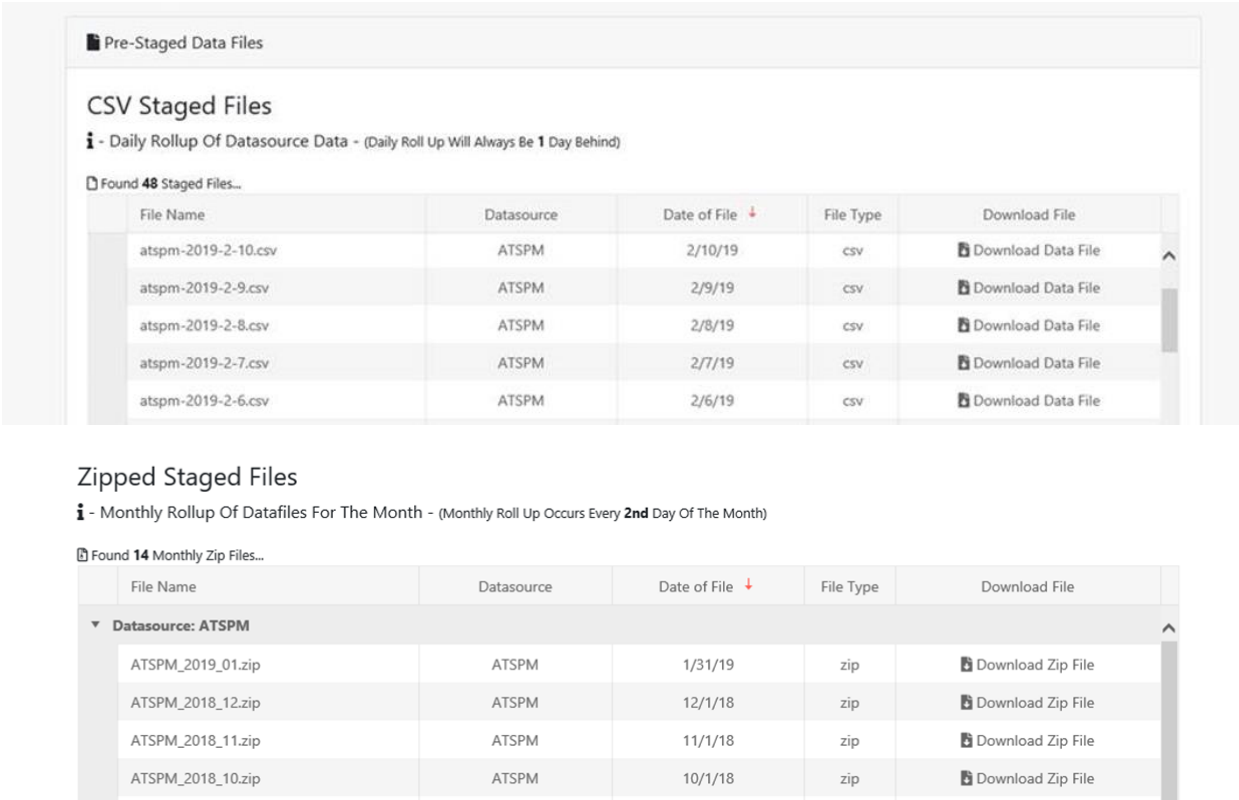
| | A | B | C | D |
|---|---|---|---|---|
| 1 | SignalID | Timestamp | EventCode | EventParam |
| 2 | 1085 | 2017-01-05 00:00:00 | 136 | 0 |
| 3 | 1085 | 2017-01-05 00:00:00 | 140 | 0 |
| 4 | 1085 | 2017-01-05 00:00:00 | 142 | 0 |
| 5 | 1085 | 2017-01-05 00:00:00 | 143 | 0 |
| 6 | 1085 | 2017-01-05 00:00:00 | 144 | 0 |
| 7 | 1085 | 2017-01-05 00:00:00 | 145 | 0 |

Figure 3: ATSPM Raw Data File

**Data Logging Requirements File**: Contains Event Code & parameter describe what each numeric value means in a table format.

| | | | | |
|---|---|---|---|---|
| Coord Pattern Change | Pattern # (0-255) | 131 | M | |
| Cycle Length Change | Pattern # (0-255) | 132 | M | |
| Offset Length Change | Pattern # (0-255) | 133 | M | |
| Split 1 Change | New Split Time in Seconds (0-255) | 134 | M | |
| Split 2 Change | New Split Time in Seconds (0-255) | 135 | M | |
| Split 3 Change | New Split Time in Seconds (0-255) | 136 | M | |
| Split 4 Change | New Split Time in Seconds (0-255) | 137 | M | |
| Split 5 Change | New Split Time in Seconds (0-255) | 138 | M | |

Figure 4: ATSPM Data Logging Requirements File



Pre-Staged Data Files

CSV Staged Files

i - Daily Rollup Of Datasource Data - (Daily Roll Up Will Always Be **1** Day Behind)

Found **48** Staged Files...

| File Name | Datasource | Date of File | File Type | Download File |
|---|---|---|---|---|
| atspm-2019-2-10.csv | ATSPM | 2/10/19 | csv | Download Data File |
| atspm-2019-2-9.csv | ATSPM | 2/9/19 | csv | Download Data File |
| atspm-2019-2-8.csv | ATSPM | 2/8/19 | csv | Download Data File |
| atspm-2019-2-7.csv | ATSPM | 2/7/19 | csv | Download Data File |
| atspm-2019-2-6.csv | ATSPM | 2/6/19 | csv | Download Data File |

Zipped Staged Files

i - Monthly Rollup Of Datafiles For The Month - (Monthly Roll Up Occurs Every **2nd** Day Of The Month)

Found **14** Monthly Zip Files...

| File Name | Datasource | Date of File | File Type | Download File |
|---|---|---|---|---|
| **Datasource: ATSPM** | | | | |
| ATSPM_2019_01.zip | ATSPM | 1/31/19 | zip | Download Zip File |
| ATSPM_2018_12.zip | ATSPM | 12/1/18 | zip | Download Zip File |
| ATSPM_2018_11.zip | ATSPM | 11/1/18 | zip | Download Zip File |
| ATSPM_2018_10.zip | ATSPM | 10/1/18 | zip | Download Zip File |

Figure 5: Daily Monthly Rollup of files in TRM

The above information in conjunction with the information described in Figure 1 and can be effectively used to estimate demand profiles and signal timing information. We also have access to the newly developed Token and Role Manager System (Figure 5). It stages the controller log files and provides access to both daily and monthly rollup of the controller log files.

### 4.2 HERE.com Data

HERE.com [4] is a company that provides mapping and location services for road traffic based in Europe. HERE.com captures both static content such as road networks etc. as well as dynamic content such as traffic flows. HERE.com collects average speed of a part of the road network (called TMC, based on OpenStreetMap designations of road portions) by using streaming data from vehicles using HERE.com GPS and routing services. By aggregating this data from various vehicles, it is able to calculate an average speed for that link at a minute-by-minute resolution.

HERE.com data set consists of data for Seminole County, Greater Orlando Metropolitan Area and tracks 1009 TMCs. Data since 2016 is available and can be queried for, by a web API. The dataset consists of the following files.

**TMC Identification File**: This file has details about the various TMCs including their OpenStreetMap identification code, their start and end positions. This file helps in identifying the location of the TMC.

| tmc | road | direction | intersection | state | county | zip | start_latitude | start_longitude | end_latitude | end_longitude | miles | road_order | timezone_name | type | country |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 102+21238 | AIRPORT BLVD | EASTBOUND | US-17/US-92/S ORLANDO DR | FL | SEMINOLE | 32773 | 28.7797 | -81.29693 | 28.7708 | -81.28328 | 1.096651 | 5 | America/New_York | P1.11 | USA |
| 102P21238 | AIRPORT BLVD | EASTBOUND | US-17/US-92/S ORLANDO DR | FL | SEMINOLE | 32773 | 28.7708 | -81.28328 | 28.76925 | -81.28036 | 0.223272 | 6 | America/New_York | P1.11 | USA |
| 102-21238 | AIRPORT BLVD | WESTBOUND | US-17/US-92/S ORLANDO DR | FL | SEMINOLE | 32773 | 28.7693 | -81.27641 | 28.76925 | -81.28036 | 0.239936 | 6 | America/New_York | P1.11 | USA |
| 102N21238 | AIRPORT BLVD | WESTBOUND | US-17/US-92/S ORLANDO DR | FL | SEMINOLE | 32773 | 28.76925 | -81.28036 | 28.76933 | -81.28096 | 0.038226 | 7 | America/New_York | P1.11 | USA |
| 102+55259 | BEAR LAKE RD | NORTHBOUND | BUNNELL RD | FL | SEMINOLE | 32703 | 28.64012 | -81.44359 | 28.65283 | -81.44376 | 0.927965 | 2 | America/New_York | P1.11 | USA |
| 102+55260 | BEAR LAKE RD | NORTHBOUND | HOLLIDAY AVE/MCNEIL RD | FL | SEMINOLE | 32703 | 28.65283 | -81.44376 | 28.65512 | -81.4438 | 0.15774 | 3 | America/New_York | P1.11 | USA |

Figure 6: TMC Identification File

**Seminole County Data File**: This file contains the actual data. For every TMC, at a minute resolution, the average recorded speed is reported. The reference speed is the desired speed that should be expected at the link (and not the maximum speed). Travel time indicates how long the vehicles took to cross the link. Confidence indicates the quality of data collection.

| tmc_code | measurement_tstamp | speed | reference_speed | travel_time_seconds | confidence |
|---|---|---|---|---|---|
| 102P10960 | 2018-09-03 00:00:00 | 31 | 22 | 36.99 | 0.69 |
| 102P10960 | 2018-09-03 00:01:00 | 31 | 22 | 36.99 | 0.69 |
| 102P10960 | 2018-09-03 00:02:00 | 31 | 22 | 36.99 | 0.69 |
| 102P10960 | 2018-09-03 00:03:00 | 31 | 22 | 36.99 | 0.69 |
| 102P10960 | 2018-09-03 00:04:00 | 31 | 22 | 36.99 | 0.69 |
| 102P10960 | 2018-09-03 00:05:00 | 31 | 22 | 36.99 | 0.69 |
| 102P10960 | 2018-09-03 00:06:00 | 31 | 22 | 36.99 | 0.69 |

Figure 7: Seminole County Data File

## 5. Conclusions

In this report, we briefly discuss the various data sources (ATSPM and HERE.com) for 300+ intersections from Seminole County that we procured for this project. ATSPM and HERE.com procured can be used for deriving

- Roadway Geometry and Corridors of Interest
- Demand profiles (volumes, turning movements, etc.) for different “base conditions” (time of day periods, days of the week, and seasons/special events)
- Signal timing parameters (green times, phasing, offset times, etc.) for the same “base conditions”

## References


[1] Zhao, Dongbin, Yujie Dai, and Zhen Zhang. "Computational intelligence in urban traffic signal control: A survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42.4 (2012): 485-494.

[2] PTV Vissim, vision-traffic.ptvgroup.com/en-us/products/ptv-vissim.

[3] “Automated Traffic Signal Performance Measures (ATSPMs).” *U.S. Department of Transportation/Federal Highway Administration*, www.fhwa.dot.gov/innovation/everydaycounts/edc_4/atspm.cfm.

[4] “HERE Technologies.” *HERE*, HERE.com.