Optimal Traffic Light Signaling Based on Genetic Algorithm Approach

   Alok    Yadav                 58090001

   Tan             Tianyu    58090005

Sajeerat Aussavaruengsuwat 58090036

Submitted to the Software Engineering Program of the International College  
in partial fulfillment of the requirements for   
Team Software Project

at the

KING MONGKUT’S INSTITUTE OF TECHNOLOGY LADKRABANG

October 2018

COPYRIGHT 2018

INTERNATIONAL COLLEGE

KING MONGKUT’S INSTITUTE OF TECHNOLOGY LADKRABANG

Table of Contents

[Abstract 1](#_Toc527062554)

[Introduction 1](#_Toc527062555)

[Chapter 1: Objectives and Problem Description 2](#_Toc527062556)

[1.1 Problem statement 2](#_Toc527062557)

[1.2 Problem Description 2](#_Toc527062558)

[1.3 Objectives 2](#_Toc527062559)

[Chapter 2: Literature Survey 4](#_Toc527062560)

[Chapter 4: Approach 5](#_Toc527062561)

[Our approach consists of three phases: 5](#_Toc527062562)

[4.1 Setting up traffic simulation model: 5](#_Toc527062563)

[4.2 Long term genetic algorithm (GA1): 5](#_Toc527062564)

[4.3 Short term genetic algorithm (GA2): 7](#_Toc527062565)

[References: 8](#_Toc527062566)

# Abstract

This report presents a methodology for overcoming the computational complexity issue of using genetic algorithms to optimize traffic signal timings over long periods of time. The methodology relies on formulating a long term traffic control strategy by using a genetic algorithm and traffic simulation model to compute fitness and then uses another genetic algorithm for short term adaptations to the strategy.

Keywords: Genetic algorithm, traffic simulation, traffic congestion, computational complexity

# Introduction

Traffic congestion is a globally occurring problem that results in massive wastes of time. More than often, the congestion is repetitive and occurs as a result of the increase in inflow from the certain sources. Current traffic management methods are far from efficient. As of right now, in Bangkok, the traffic signal timings are constant. They don’t change depending on the traffic conditions. Managing traffic signal timings based on traffic conditions is an efficient way of reducing traffic congestion. Genetic algorithms can be used to optimize traffic signaling based on changing traffic conditions and have been shown to be more efficient compared to other evolutionary strategies [6].

There have been many researches that attempt to use genetic algorithms to optimize traffic signal timings. However, they mostly focused on real time control based on existing traffic conditions. Due to the time constraints of real time control, the simulation time for most of the researches was only of the order of a few minutes. Actions on traffic signals can have long term consequences that are overlooked when using short time simulations. Furthermore, this approach fails to capitalize on the history of traffic, which can help in managing traffic more efficiently. The difficulty of optimizing traffic signaling over a long period of time arises due to the computational difficulty of running repeated simulations over long durations of time over a large area. It has been shown that expanding the area considered yields better optimization results [4], and increasing simulation time will allow the genetic algorithm to find better solutions. However, the potential of using genetic algorithms for traffic optimization hasn’t fully been realized due to the time complexity issues.

In our project, we explore two ways of combating the high computation time. Firstly, we can use a macroscopic traffic model to simulate traffic, in order to reduce computation time. Macroscopic models aggregate the description of traffic flow; macroscopic models use speed, flow and density as measures of effectiveness [2]. In the second approach, instead of having a single genetic algorithm to optimize traffic over a long period of time, we try to use a long term genetic algorithm to optimize recurring traffic congestions. The long term genetic algorithm will give us the optimal traffic density points for the roads throughout the control period. The second genetic algorithm will try to change traffic signaling to match the density points produced by the long term genetic algorithm. By this approach, we can plan over a long period of time, thereby improving rush hour traffic. Furthermore, our approach incorporates the history of traffic; hence it should be better suited at managing recurring traffic congestions.

# Chapter 1: Objectives and Problem Description

## **1.1 Problem statement**

Genetic algorithms can be used to efficiently optimize traffic signals and reduce traffic congestion; however the time complexity of the task is too high to be practical.

## **1.2 Problem Description**

One of the most cost effective ways of managing traffic is optimizing traffic signal timings. Genetic algorithms can be used to find near optimal traffic signal timings. However, prior researches on the application of genetic algorithms to the traffic setting problem have had limited results due to the computational complexity of simulating traffic. As a result, optimizing traffic over long periods of time has proven to be problematic. This project aims to overcome this obstacle by performing the computations for long term traffic control prior to real time control and formulating a traffic control strategy. Then using a real time algorithm to adapt to the pre computed long term strategy. The application of this method is expected to reduce the total waiting time of cars by more than 10% for the area considered in [4].

## **1.3 Objectives**

As of right now, we have established the following major objectives:

**Gather travel speed data for roads in the selected region:**

This data is going to be used for the long term genetic algorithm, as well as to understand the variability of traffic, which will help us improve the short term genetic algorithm.

As of right now, we have implemented and tested the scripts to collect traffic data using TomTom API. We still have to test the scripts for converting travel speed data to density data and for using density data to determine the flow of traffic at intersections.

**Set up traffic simulation model:**

In order to use evolutionary algorithms, we need a mechanism to predict the fitness of a solution. We are going to need a traffic simulation model that can accurately predict traffic conditions based on changing traffic inflow and changing traffic signaling.

Thus far, we have acquired permission to use the traffic simulation framework developed by Pawel Gora. Traffic simulation framework is an advanced tool for simulating and investigating real vehicular traffic in cities [8]. In the event that the time complexity of the task prevents the usage of traffic simulation framework, we will have to look into developing a mesoscopic traffic simulation model.

**Implement and experiment with long and short term genetic algorithms:**

In this research, we will have to experiment with the simulation parameters and different variations of genetic algorithms, to determine the best way to optimize traffic congestion. The details about the function and implementation about the genetic algorithms are given in the following section.

# Chapter 2: Literature Survey

The problem of optimizing traffic flow using traffic light signaling is a complex one and has been researched often. Many methods have been tested in their capability of optimizing traffic. These Methods include reinforcement learning, genetic algorithms, swarm algorithms, neural networks, organic computing and fuzzy logic [3]. For our project, we focused mostly on genetic algorithms. Most of the papers researched in this literature review focus on the application of genetic algorithms and different traffic simulation methods to the traffic optimization problem.

[3]: In this paper, Pawel Gora uses a genetic algorithm to optimize the flow of traffic using traffic simulation framework. The major shortcoming of the approach was the low simulation time of 600 seconds. This was due to the computational complexity of the simulations that were microscopic in nature. Microscopic models are models that continuously or discretely predict the state of individual vehicles [2]. As a result, the improvement in traffic conditions was only minor (3.1%).

[4]: This paper builds upon the work of Pawel Gora. It used a modified version of the traffic simulation framework and a high performance computing cluster to overcome the computational limitations. The results obtained were slightly better than in [3]. However, the results were still not satisfactory, in spite of running much more iterations of algorithms (50 populations). The results obtained after using a mesoscopic traffic simulation were much better. A significant result of this research was the impact of the area simulated and the area used for computing fitness. Performance is improved by expanding area optimized and reducing area for computing fitness.

[5]: This paper also focuses on real time control of traffic by using genetic algorithms to optimize traffic in a microsimulation. The scale of the simulation however, is smaller. The simulation model consisted of only six intersections with one way roads. The results from the prior paper show that optimization will not yield successful results when considering a small area; this is reflected by the fluctuations in the graph of performance and generation number in this paper.

[6]: A graph model is used to represent traffic in this research. They devise a branch and bound algorithm to obtain the optimal solution [6], this method takes a very long time in case of large graphs though. The paper also explores the effects of using other evolutionary algorithms such as genetic algorithms, particle swarm optimization and ant colony optimization. Amongst the tested evolutionary algorithms, genetic algorithms had the best performance.

[7]: In this paper, a genetic algorithm is used for real time optimization. But the algorithm also takes into account the importance of a road in the intersection. The parameter optimized is the total number of cars on a road. The results compared the efficiency of the genetic algorithm in comparison to a fixed timing system. The improvement was of almost 22%. That seems too good, in comparison to prior papers. We suspect that this improvement is due to the different models; similar to the improvement of performance in [4] when a mesoscopic model was used instead of a microscopic one.

The reviewed papers tend to focus on real time control of traffic based on the current traffic conditions. There are two shortcomings of this approach. Firstly, traffic congestions have a tendency to repeat, optimizing based only on current traffic conditions fails to capitalize on this important property of traffic. Secondly, actions on traffic have long term consequences that real time control will fail to consider. Real time control forces the use of short term simulations due to computational cost of long term simulations. Hence, we propose a new strategy for traffic optimization that formulates a long term strategy and focuses on short term adaptation to the strategy.

# Chapter 3: Background Knowledge

# This project associates with the following fields of research:

# Traffic simulation

# Traffic Simulation Framework

# Evolutionary algorithms

# 3.1 Traffic Simulation

# Traffic simulation models can be divided into three main categories:

# Microscopic:

# Mesoscopic:

# Megascopic:

# The traffic models can be compared based on their accuracy to real life traffic and the computational time of the simulations.

# Chapter 4: Approach

# Our approach consists of three phases:

* Setting up a traffic simulations that encapsulate the day to day variation in traffic
* Using a long term genetic algorithm to optimize recurring traffic congestion over a long period of time.
* Use a short term genetic algorithm to try to mimic the density points outlined by the long term genetic algorithm, for traffic conditions similar to the one optimized by the long term genetic algorithm.

## **4.1 Setting up traffic simulation model:**

The traffic simulation model is meant to provide predictions of traffic conditions based on the changing traffic signal timings and changing traffic flow. We are going to use a series of simulations with similar traffic conditions, to test how our hypothesis about using past simulations (GA1) to better improve similar traffic conditions at a later stage (GA2).

In order to set up the model, we need to do the following:

* Setup a model that can accurately predict traffic based on the changing inflow and outflow of traffic from each road.
* Select a region to model the traffic
* Acquire data about the inflow and outflow of traffic for each road during the considered time period. To get information about traffic flow, we need to do the following:
  + Acquire Travel speed data at each road near the intersections.
  + Convert travel speed data into approximate density data.
  + Calculate the fraction of cars exiting into each road from an intersection using the density data

## **4.2 Long term genetic algorithm (GA1):**

Once we have the traffic data over a time period that relates to repeating traffic congestion, the long term genetic algorithm is meant to optimize the traffic signals over the entire period. The output of this process is going to be the optimal traffic densities at each road after small time intervals. This task is computationally expensive and will have to be done prior to the real time control task.

**Solution Domain:**

In order to describe the process of obtaining the best solution, we need to define the solution space. The following definitions describe the structure and significance of a genotype for the long term genetic algorithm.

**Definition 1:** *Let be the set of traffic lights at a single crossroad.* ***Representant*** *of the set A is any element of the set A. It will be marked as .* ***Representant*** *of any element is . The choice of representant is important, because different representants will yield differing signal timings, however, the eventual result will be the same change in fitness, even though the genotype might look different.*

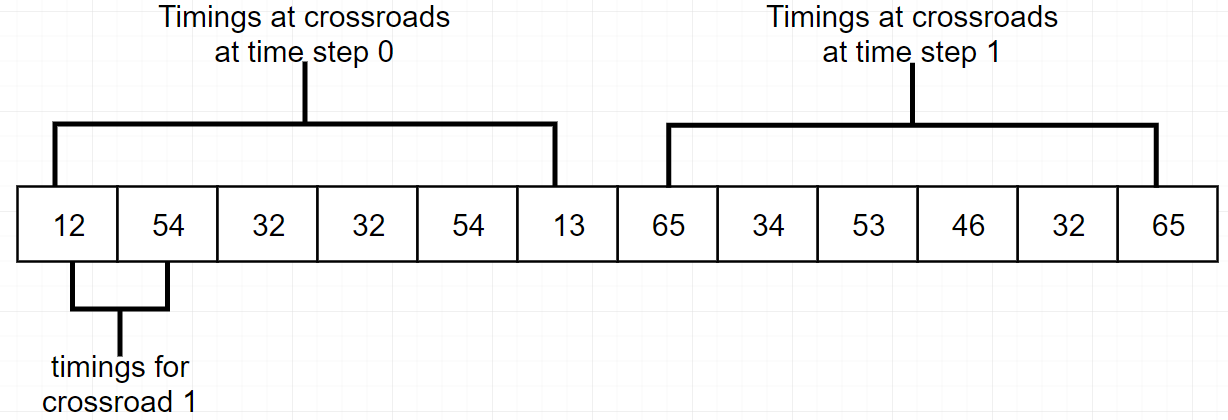
**Definition 2:** *Let be the set of all crossroads in the road network. Let be the set of* ***representants of all crossroads****.*

**Definition 3:** *Let be the set of* ***possible phase durations*** *for the red and green traffic signals, where correspond to the maximum and minimum phase durations of red and green signals respectively.*

**Definition 4:** *Let be the set of time steps. Let g: G → N be any function mapping the set of representants to the set of possible phase durations and let be a set of different possible mappings from the set of representants to the possible phases.* ***Genotype*** *for the road network is any one to one mapping .*

The above definitions are based on genotype definitions in [3]. A genotype represents the signal phases for each intersection of the road network, throughout the time steps. Initially, the genotypes will be assigned randomly (random phases). Then the genotypes will be used for evolution. A gene in the genotype represents the timing of either the red or green signal at some crossroad at some point in time.

Example: If there were 3 crossroads in the area to be optimized, and the genetic algorithm was considering optimization over two time steps, a single genotype would look like this:



**Fitness function**:

For the long term genetic algorithm, we are going to base the fitness of a genotype on either of the following attributes:

* Peak car density at any road in the road network.
* Total waiting time for cars.
* Average travel speed.

We can define and test many fitness functions using the traffic simulation framework; this is not the final list of fitness functions. The primary measure of fitness however is going to be the total waiting time for cars, based on the suggestions of Pawel Gora.

**Selection:**

If there are N individuals in the population, we will take of the individuals with the best fitness score.

**Crossover:**

Each selected individual will be crossed with every other selected individual. When two individuals are crossed, the new individual will have genes randomly selected from either of the parents, with equal probability.

**Mutation:**

Each gene in the newly created individuals will be randomly changed to some random value from the set of possible values for signal timings using probabilities from the set {0.001, 0.01, 0.1} in successive experiments (different probability in different experiments).

**Termination:**

We will not terminate the genetic algorithm based on the fitness of the individuals, but based on the computation time; we are yet to determine the time limit.

**Simulation Specifications:**

The specifics for the simulation that the long term genetic algorithm will try to optimize are as follows:

* Time duration: 60 minutes (we will increase the duration later)
* Time step duration: 5 minutes
* Time steps (n): (Time duration)/(Time step duration) = 12

**Output of GA1:**

**Definition 5:** *Let D be the set of densities at each road in the network, , where is the density at road k. Let denote the set of road densities at each road in the network at time step t for the best genotype obtained from the long term genetic algorithm. Let DTT be the set of densities at each road in the network at each time step . Then the* ***output*** *of the algorithm is DTT.*

## **4.3 Short term genetic algorithm (GA2):**

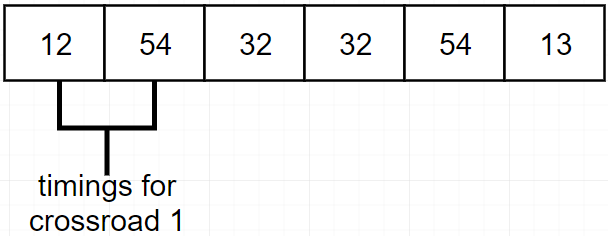
The short term genetic algorithm will be used to achieve real time control of traffic. Instead of trying to find the traffic signal timings to get optimal traffic densities like in GA1, this genetic algorithm focuses on finding traffic signal timings that will reproduce the traffic densities produced by the long term genetic algorithm. Hence the performance measure for this genetic algorithm will be the difference between the simulated traffic densities produced by the individuals in the population and DTT.

**Solution domain:**

**Definition 6:** *Let be the set of all crossroads in the road network. Let be the set of representants of all crossroads. Let be the set of possible phase durations for the traffic signals, where correspond to the maximum and minimum durations of red and green phases respectively.* ***Genotype*** *for the road network is any function or.*

The above definition is based on genotype definitions in [3]. A genotype represents the signal phases for each intersection of the road network. Initially, the genotypes will be assigned values based on the evolved genotypes from GA1 corresponding to the same time step. Then the genotypes will be used for evolution. A gene in the genotype represents the timing of either the red or green signal at some crossroad.

Example: If there were 3 crossroads in the area to be optimized, a single genotype would look like this:



**Fitness function**:

In order to specify the fitness of a genotype, we need to define the attribute to be used to measure the performance. We use the traffic density as the attribute to measure fitness. The fitness of a genotype is defined as follows

**Definition 7:** *Consider a genotype G generated by GA2 corresponding to time step i from the long term simulations conducted prior and its output densities D for each road, , where is the density at road k. The* ***fitness of G*** *is given by the equation , where DTT is the optimal densities at each time step obtained from the long term genetic algorithm, are the optimal densities at each road at time step i.*

Apart from the above defined measure of fitness, we will also experiment with the fitness functions defined in section 4.2

**Selection:**

If there are N individuals in the population, we will take of the individuals with the best fitness score.

**Crossover:**

Each selected individual will be crossed with every other selected individual. When two individuals are crossed, the new individual will have genes randomly selected from either of the parents, with equal probability.

**Mutation:**

Each gene in the newly created individuals will be randomly changed to some random value from the set of possible values for signal timings using probabilities from the set {0.001, 0.01, 0.1} in successive experiments (different probability in different experiments).

**Termination:**

We will not terminate the genetic algorithm based on the fitness of the individuals, but based on the computation time; we are yet to determine the time limit.

**Output:**

The output for the short term genetic algorithm is going to be the traffic signal timings to make the traffic as similar as possible to the traffic simulated in the long term genetic algorithm.

# Chapter 5: Development/Implementation

# In this section, we describe the experiments conducted and the implementation of the systems required to conduct the experiments.

# This section can be broken down into the following parts:

# Experiments

# Data gathering

# Genetic Algorithms

# GUI application

# 5.1 Experiments

# In our research, we are going to experiment with the following parameters:

# Variability between simulations used by GA1 and GA2

# Time step duration

# Number of time steps considered by GA1

# Fitness functions for GA1 and GA2

# Area whose traffic signals are optimized by genetic algorithms

# Area used to calculate fitness: If this area is smaller than the area used to optimize traffic signals, the fitness of the solutions should be higher. We expect this because of the results given by [4]

# Effect of optimization of a fitness function on other fitness functions

# We define the following specifications for the parameters to be experimented with:

# Fitness function1: Time spent travelling under 20km/h, can be used by both GA1 and GA2

# Fitness function2: Time spent not moving, can be used by both GA1 and GA2

# Fitness function3: Difference between densities at roads between solutions obtained by GA1 and GA2 at the same time step. Used only by GA2, specifics in section 4.3

# Fitness function4: weighted fitness (the fitness value of a certain region computed by using fitness functions 1, 2, or 3 is multiplied by a certain weight)

# Area A:

# Area B:

# 5.1 Data gathering

# In order to gather the necessary traffic data, to set up the traffic models, we need to do the following:

# Identify and index all the traffic intersections in the selected region.

# Identify and index all the unsegmented road sections (no roads feed in or out from the chosen road).

# Obtain the number of lanes and travel speed data at each road near the traffic intersection that feeds into the road using tomtom traffic API.

# Convert all travel speeds into car densities [insert pseudo algorithm].

# Use the density data to obtain information about what proportion of cars go to which roads at what times [insert pseudo algorithm].

# 5.2 Genetic Algorithms

# The genetic algorithms were implemented using the Distributed Evolutionary Algorithms in Python (DEAP) module in python.

# 

# In order to simulate traffic settings, the genetic algorithms use the Simulator object. The simulator object uses the requestStats function to post a request to the Traffic Simulation Framework, which is a microservice in the azure cloud. Currently, we only send the traffic signal offsets for each traffic intersection in the request. The response is the fitness of the individual that corresponds to the traffic signals posted. The run time of one experiment can go up to a month, hence we had to parallelize the requests

# 5.3 GUI application

# References:

[1] N. T. Ratrout and S. M. Rahman, "A COMPARATIVE ANALYSIS OF CURRENTLY USED MICROSCOPIC AND MACROSCOPIC TRAFFIC SIMULATION SOFTWARE", *The Arabian Journal for Science and Engineering*, vol. 34, no. 1, pp. 121-133, 2009.

[2] S. A. Boxill and L. Yu, "An Evaluation of Traffic Simulation Models for Supporting ITS Development", Center for Transportation Training and Research Texas Southern University, Houston, Texas, 2000.

[3] Gora P., “A genetic algorithm approach to optimization of vehicular traffic in cities by means of configuring traffic signals”, in “*Emergent Intelligent Technologies in the Industry*”, 2011, pp. 1-10, ISBN: 978-3-642-22731-8.

[4] P. Gora and P. W. Pardel, "Application of Genetic Algorithms and High-Performance Computing to the Traffic Signal Setting Problem", 2015.

[5] Kwasnicka, H., Stanek, M.: Genetic Approach to Optimize Traffic Flow by Timing Plan Manipulation. In: Yuehui, C., Abraham, A. (eds.) *ISDA 2006 Proceedings*, vol. II, pp. 1171–1176. IEEE Computer Society, Los Alamitos (2006)

[6] Chen, S.W., Yang, C.B., Peng, Y.H.: Algorithms for the Traffic Light Setting Problem on the Graph Model. In: Proc. of *the 12th Conference on Artificial Intelligence and Applications*, TAAI (2007)

[7] L. Singh, S. Tripathi and H. Arora, "Time Optimization for Traffic Signal Control Using Genetic Algorithm", *Int. J. of Recent Trends in Engineering and Technology*, vol. 2, no. 2, 2009.

[8] Gora P., “Traffic Simulation Framework”, in “*14th International Conference on Modelling and Simulation*”, 2012, pp. 345-349, ISBN: 978-0-7695-4682-7