Traffic Signal Timings Optimization based on Genetic Algorithm and Gradient Descent

Alok Yadav
Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
alok20121998@gmail.com

Chaiwat Nuthong
Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
chaiwat.nu@kmitl.ac.th

***Abstract*—Traffic congestions are a recurring problem that results in significant losses both financially and environmentally. Optimizing traffic signal timings is one of the most cost-effective ways to mitigate such effects. Optimization of traffic signal timings capable of minimizing congestion is, however, computationally expensive. Research needs to be conducted in order to develop algorithms capable of better optimization using fewer computational resources. This paper presents a novel approach to traffic signal optimization that combines genetic algorithms and gradient descent to obtain optimized traffic signal timings. The genetic algorithm is used to arrive at a starting point for gradient descent; gradient descent is then used to obtain further improvement.**

Keywords-component; traffic signal timings; genetic algorithm; gradient descent

I. Introduction

With the growing number of operating automobiles, the demands on road infrastructure have also grown. Inability of the infrastructure to handle such demands results in traffic congestion. In order to alleviate this repercussion, either the infrastructure must be changed, or, the infrastructure must be utilized more efficiently. The prior requires considerable expenses. The latter, however, can be accomplished by better control of traffic signal phase durations, which is comparatively inexpensive. Such systems are already being used, such as SCOOT, SCATS and FHWA [1]. Optimizing traffic flow by controlling traffic signal timings does, however, require research into how traffic signals ought to be controlled to better optimize traffic flow.

Several researches have been conducted regarding traffic optimization using traffic signal timings. These researches broadly fall under two categories. Firstly, re-searches that use population based evolutionary algorithms to carry out real-time optimization; such algorithms include genetic algorithms [2-6], particle swarm optimization [5, 7] and ant colony optimization [5]. Secondly, researches that train models prior to deployment. Model training may be done via reinforcement learning [8-10] or deep learning [11-12]; other approaches might rely on fuzzy rule-based systems [13]. All approaches have in common some way of simulating traffic. For population-based approaches, this is necessary to evaluate fitness of the generated individuals. In the case of pre trained models, a simulator is needed to generate data required for training neural networks, constructing reward functions, etc. A challenge faced by all research in the field is that of traffic simulation detail. Simulations need to be detailed and should attempt to mimic reality as closely as possible for the algorithms to be applicable to real life scenarios. However, if the simulations are too detailed, training in case of pre-trained models might become infeasible in terms of the time required. For population-based optimization algorithms, if simulations become too long, the algorithm might no longer be able to meet real-time constraints. This obstacle to traffic optimization creates a need for optimization algorithms that can better utilize simulation results to carry out optimization.

This research presents a novel approach to traffic optimization. In the proposed approach, a genetic algorithm is used to arrive at a sub-optimal set traffic of traffic signal timings. A gradient descent (GD) algorithm is then used to further improve the traffic signal timings. The performance of the proposed approach is compared with that of a genetic algorithm-based approach across varying optimization periods. Simulation of Urban Mobility [14] is used to test the performance of the implemented algorithms.

II. Problem Formulation

Consider a road network with multiple traffic signal-controlled intersections; each intersection has multiple synchronized traffic signals. A single traffic signal can be used to represent all traffic signals in an intersection. Any traffic signal T has three phases (red, green and yellow). Assuming a constant sum of phase durations (120 seconds in this paper) and constant yellow phase duration, a phase offset (difference between red and green phase durations) can be used to represent all phase durations of a traffic signal. This offset for traffic signals at intersection I is represented as $P(I)$.

The traffic signal timings for a road network with n intersections $\{I_1 \ldots, I_n\}$ will be a set $\{P(I_1) \ldots, P(I_n)\}$. Each set of traffic signal timings will impact traffic differently. The objective is to find a set of traffic signal timings that minimizes the total waiting time of the cars in the road network.

III. Traffic Simulation

Traffic Simulations for the experiments were conducted using Simulation of Urban Mobility [14]. Simulation of Urban Mobility (SUMO) allows simulation of areas selected from Open Street Map. During experiments, the area between latitude range 52.2330-52.2203 and longitude range 21.0005-21.0283 was simulated. 37 traffic lights were controlled in the area to optimize traffic. The sum of phase durations for each traffic light was set to 120 seconds. During simulations, the phase offset was varied from -60 to 60 seconds (phase durations range from 30 to 90 seconds). The phase cycle is determined automatically by SUMO. The offset then is the difference between the two longest phases.

IV. Genetic Algorithm Approach

The genetic algorithm (GA) approach is a simplistic approach that serves as a baseline for comparison with the proposed approach. The steps in a single GA based optimization are described as follows:

1. GA generates a population of random genotypes, each genotype representing the timings at traffic signals across the road network for the considered time duration (optimization period)
2. The fitness of these genotypes is calculated by using the traffic simulator, to measure how well each genotype can handle traffic
3. The GA selects some of the genotypes based on their fitness and uses them to generate the next population of genotypes
4. Step 2 now repeats with new population, unless a termination criterion has been met, such as reaching the allowed number of generations

Genetic Algorithm

The components and operators of the genetic algorithm used in the genetic algorithm approach are described as follows

1) Genotype: A genotype represents a set of traffic signal timings for the road network. This section mathematically describes a genotype. The timings are represented using phase offsets. For a traffic signal, the phase offset is the difference between the two largest phases, i.e. the red phase and green phase. The order of the phases is determined by SUMO. The set of possible phase offsets is denoted by *N*.

$$N = \{min_p, \ldots, max_p\} \tag{1}$$

min_p and max_p represent the minimum and maximum possible phase offsets allowed. In this paper, min_p is -60 and max_p is 60. Consider an intersection *I* with a few traffic signals. These traffic signals are going to be synchronized, i.e. they will have the same phase durations that are offset so that when one signal has a phase, another will have the opposing phase. Hence, the phase offsets of all traffic signals at an intersection can be represented by a single variable. This variable is denoted as *P(I)*. The set of all intersections in the road network is denoted by *C*

$$C = \{I_1, I_2, I_3, \ldots, I_n\} \tag{2}$$

All the phase offsets in the network can be represented by a set *G*

$$G = \{P(I_1), P(I_2), P(I_3), \ldots, P(I_n)\} \tag{3}$$

Genotype for the road network is any function from G to N. The genotypes used in the genetic algorithm will be initialized randomly.

2) Fitness Function: The fitness of a genotype is the total waiting time of simulated cars in seconds measured by the traffic simulator when simulating the traffic signal timings specified by the genotype.

3) Selection: From a population of size *S*, the genotypes with the highest fitness are selected. The number of genotypes selected is equal to the sqare root of *S/2*.

4) Crossover: Each selected genotype is crossed with every other selected genotype using two-point crossover. Each crossover takes two genotypes and produces two genotypes.

5) Mutation: A gene, which represents the timings for a single crossroad, may be randomly switched with another gene in the genotype based on the mutation probability

V. Proposed GA-GD Approach

The task in optimizing traffic signal timings to reduce waiting time is to find a point in the search space of traffic signal timings that corresponds to a waiting time lower than other points. Upon a few executions of the above-mentioned approaches, it becomes apparent that genetic algorithms are capable of converging in the general vicinity of an optimal solution but have trouble converging to the actual optima. This is because the genetic algorithm approach relies too much on randomness. Gradient descent on the other hand relies on making small changes in the direction that offers the best improvement in fitness. For this reason, the proposed GA-GD approach uses a variation gradient descent to fine tune the individual created by the genetic algorithm.

In the proposed GA-GD approach, the genetic algorithm is used to arrive at a sub-optimal solution. Then a gradient descent-based algorithm is used to fine tune the solution to improve its fitness. The genetic algorithm component of the proposed GA-GD approach is identical to that of the genetic algorithm approach. The gradient descent component is described as follows.

Gradient descent is carried out over multiple iterations, each with a progressively smaller step size used to increment or decrement the genes. During an iteration, the gradient vector is calculated by moving the provided genotype a step forward and backward in every possible direction and calculating the fitness of the resulting genotype. The gradient vector hence consists of twice as many elements as there are genes in the genotype, each element representing the impact of increasing or decreasing a gene. In gradient descent, usually the genotype would be moved slightly in every direction based on the gradient vector. Here, however, moving the genotype in every direction results in no improvement in fitness. The genotype is hence first moved in the direction that offers the most improvement in fitness. Then it is moved in the direction that offers the second-best improvement. If the new move does not result in an improvement in fitness, the move is reversed. This is continued with the third best move, fourth best move and so forth. There is a limit, however, to the number of moves that will be considered; referred to as (n_{step}). This is because only the first few moves offer improvement in fitness.

A single iteration of gradient descent consists of two steps. Firstly, a list of variations that give the best improvement in fitness is assembled; this list refers to the first list of genotype variants in Figure 1. From this list, n_{step} variations with the lowest fitness are selected to get the second list of genotype variants. Next, each of the n_{step} variations are successively applied to the original genotype, to compound the improvement of the genotype. This is illustrated in Figure 2.

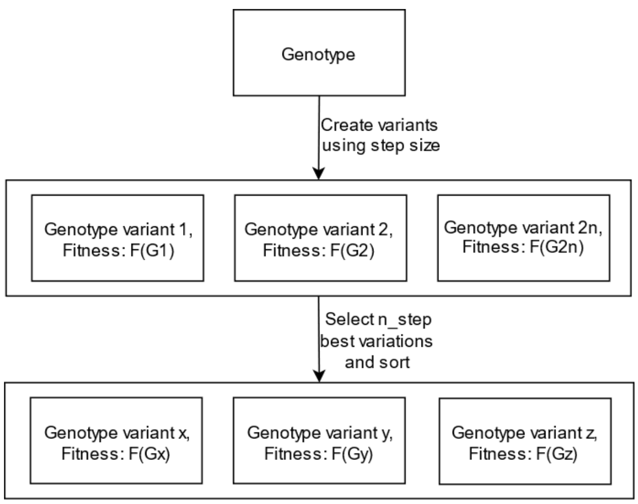

Figure 1. A figure illustrating how n_{step} variations to be considered are selected.

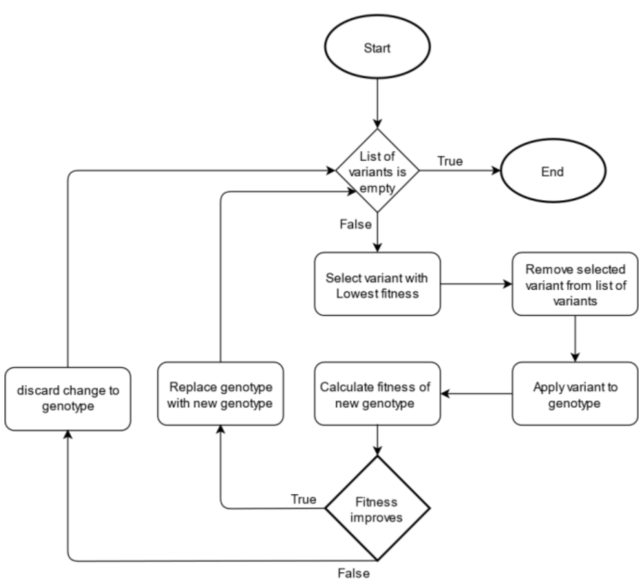

Figure 2. A flowchart displaying how the variations are selected to be applied to the genotype.

VI. Experiments

Experiments were conducted to compare the performance of the GA approach with the GA-GD approach. A single experiment consists of executing a genetic algorithm for how many ever generations are required by the GA-GD approach, the last generation is then used separately by the GA and GA-GD approaches. To obtain the GA approach performance, the genetic algorithm continues from the generation at which the GA-GD approach forked out. To obtain the GA-GD approach performance gradient descent is performed on the best genotype from the generation at which the GA-GD approach forked out, as described in the previous section.

Experiments were carried across varying levels of traffic and varying optimization durations. The level of traffic is controlled by changing the arrival period (p) at which new vehicles are generated. Optimization duration refers to the duration of the simulations that are used during optimization. The performance of an approach is measured in terms of the total waiting time of simulated cars when the simulation traffic lights follow the timings dictated by the optimization algorithm.

A. *Experiment Setup*

During the experiments, optimization durations of values 4 min, 6 min and 8 min were considered. The value of the arrival period (p) was varied from 0.3 to 0.5. Specifications for the algorithms used in each approach are listed in Table 1.

n_{step} was set to 10 as this results in a gradient descent iteration taking the same amount of computation as a genetic algorithm generation with population size 84.

The experiments were conducted on an ASUS ROG G751JT notebook; the following are its specifications:

5. Processor: Intel® Core™ i7 4710HQ Processor
6. OS: Windows 10 home
7. Chipset: Intel® HM87 Express Chipset
8. Memory: DDR3L MHz SDRAM, 16GB
9. Graphic: NVIDIA® GeForce® GTX970M with 3GB GDDR5

TABLE I. Parameters used for each approach during experiments

Parameter	**GA**	**GA-GD**
GA population size	84	84
GA number of generations	5	3
GA mutation probability	0.1	0.1
No. of GD iterations	-	2
GD step size	-	30
n_{step}	-	10

B. Results and Discussion

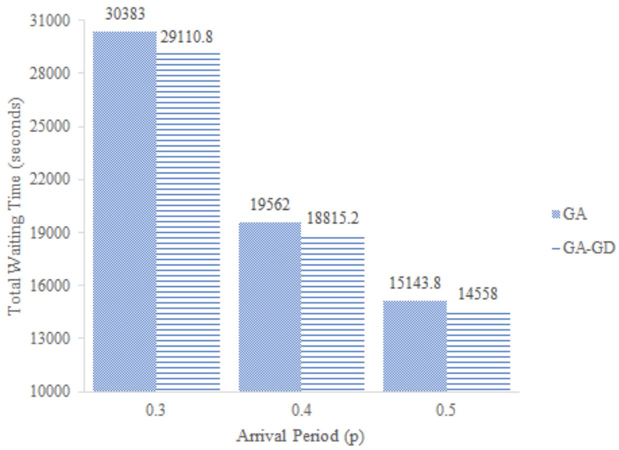

Figure 3. Performance when optimization duration is 4 min.

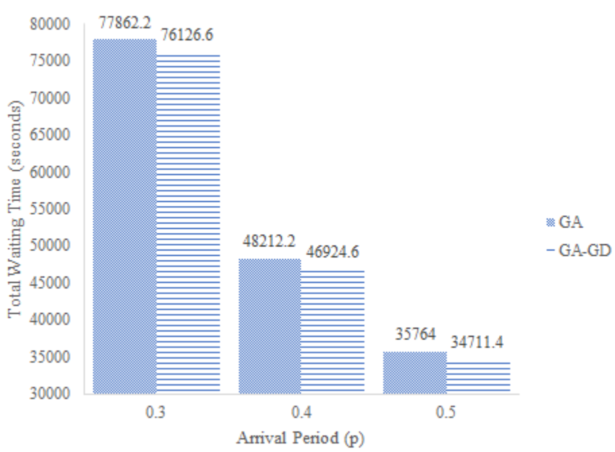

Figure 4. Performance when optimization duration is 6 min

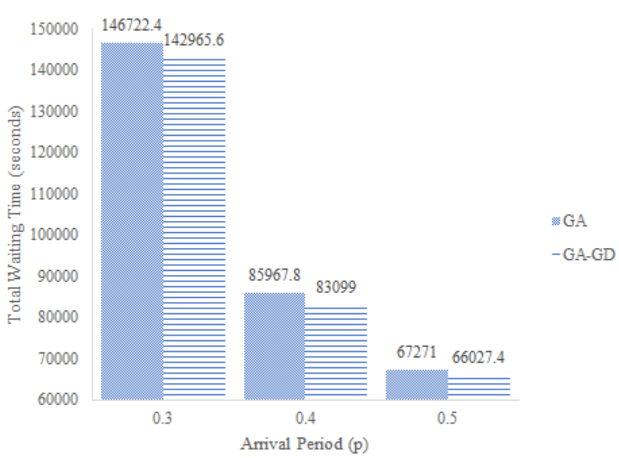

Figure 5. Performance when optimization dration is 8 min.

The GA-GD approach is able to outperform the GA approach in all experiments by a varying degree. The waiting time is decreased in most cases by more than 2% and up to 4% in a few cases. Due to the stochastic nature of the proposed approach, there were multiple experiments where GD failed to improve upon the genotype. Performing GD on multiple genotypes may alleviate this issue.

VII. Conclusion and Future Work

This paper presents a novel approach to traffic optimization that combines genetic algorithm and gradient descent. The proposed approach is able to outperform the GA based approach consistently by a significant margin. Using GD reduces the waiting time by thousands of seconds within a few minutes. When applied across large areas for long periods of time, such a reduction in waiting time can significantly reduce pollution and improve travel times.

Further research remains to be conducted on using gradient descent to optimize traffic signal timings. Improvements can be made in how the learning rate changes between iterations. Experiments should be conducted with multiple starting points instead of one, in order to achieve better and more consistent convergence. Additionally, the heuristic that the algorithm is based on itself may be sub-optimal and can be experimented with to create algorithms that offer better improvement in traffic and require less computation.

References

[1] K. Fehon and J. Peters, “Adaptive traffic signals, comparison and case studies,” In ITE District 6 2004 Annual Meeting, 2004.

[2] P. Gora, “A genetic algorithm approach to optimization of vehicular traffic in cities by means of configuring traffic signals,” In Emerging Intelligent Technologies in Industry, Studies in Computational Intelligence, vol 369, pp, 1-10. Springer, Berlin, Heidelberg, 2011.

[3] P. Gora and P.W. Pardel, “Application of genetic algorithms and high-performance computing to the traffic signal setting problem,” In Concurrency, Specification and Programming, 2015.

[4] H. Kwasnicka and M. Stanek, “Genetic approach to optimize traffic flow by timing plan manipulation,” In Sixth International Conference on Intelligent Systems Design and Applications, vol. II, pp. 1171–1176, Jinan, 2006.

[5] S. Chen, C. Yang and Y. Peng, “Algorithms for the traffic light setting problem on the graph model,” In Taiwanese Association for Artificial Intelligence, pp. 1-8, 2007.

[6] J. Javier, Sanchez-Medina, J. Manuel and R. Enrique, “Traffic Signal Optimization in “La Almozara” District in Saragossa Under Congestion Conditions, Using Genetic Algorithms, Traffic Microsimulation, and Cluster Computing.” In IEEE Transactions on Intelligent Transportation Systems, vol. 11, pp. 132-141, 2010.

[7] Y. Qian, C. Wang, H. Wang and Z. Wang, “The optimization design of urban traffic signal control based on three swarms cooperative-particle swarm optimization,” In 2007 IEEE International Conference on Automation and Logistics, pp. 512-515, Jinan, 2007.

[8] L. Kuyer, S. Whiteson, B. Bakker and N. Vlassis, “Multiagent reinforcement learning for urban traffic control using coordination graphs,” In Machine Learning and Knowledge Discovery in Databases, vol 5211, Springer, Berlin, Heidelberg, 2008.

[9] M. Coşkun, A. Baggag and S. Chawla, “Deep reinforcement learning for traffic light optimization,” In 2018 IEEE International Conference

on Data Mining Workshops (ICDMW), pp. 564-571, Singapore, 2018.

[10] X. Liang, X. Du, G. Wang and Z. Han, “A deep reinforcement Learning network for traffic light cycle control,” In IEEE Transactions on Vehicular Technology, vol. 68, no. 2, pp. 1243-1253, 2019.

[11] L. Jiang, Y. Li, Y. Liu and C. Chen, “Traffic signal light control model based on evolutionary programming algorithm optimization BP neural network,” In 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), pp. 564-567, Macau, 2017.

[12] Z. Liu, “A survey of intelligence methods in urban traffic signal control,” In International Journal of Computer Science and Network Security, vol. 7, No. 7, 2007.

[13] S. Chiu and S. Chand, “Adaptive traffic signal control using fuzzy logic,” In Proceedings. The First IEEE Regional Conference on Aerospace Control Systems, pp. 122-126, Westlake Village, CA, USA, 1993.

[14] M. Behrisch, L. Bieker, J. Erdmann and D. Krajzewicz, “SUMO – Simulation of Urban MObility: An Overview, In Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation, ThinkMind, Barcelona, 2011.

Authors’ Background

Name	Email	Position (Prof , Assoc. Prof. etc.)	Research Field	Homepage URL
Alok Yadav	**alok20121998@gmail.com**	**Student**	**Computational Intelligence**	
Chaiwat Nuthong	**chaiwat.nu@kmitl.ac.th**	**Assistant Professor**	**Computational Intelligence**	