

Application of a genetic algorithm to reduce the search time for traffic signal control

H. Kazama¹, H. Mochizuki², S. Takahashi² & H. Nakamura²

¹*Kyosan Electric Manufacturing Co., Ltd., Japan*

²*College of Science & Technology, Nihon University, Japan*

Abstract

In order to minimize average travel time (ATT) through traffic signals, a traffic parameter offset and optimized offset pattern are discussed. For practical use, the search time of the optimized offset pattern should be as short as possible. To this end, a genetic algorithm is introduced. In the conventional method of searching for an optimal solution, a target fitness value is not given, making it extremely difficult to reduce the search time. To overcome this problem, here we introduce a new method that depends on the ATT distribution characteristics and that derives the target fitness value by using the target road parameters. In addition, the effectiveness of the method is illustrated with a case study.

Keywords: traffic signal, genetic algorithm, offset, search time, average travel time.

1 Introduction

One method of searching for an optimal solution is a genetic algorithm (GA). If a GA is given a target value of the optimal solution, it can be terminated upon reaching that solution; however, if the target value is unknown, a method such as treating the convergence conditions of the solution search as a basis for terminating the algorithm, or treating the best solution obtained during several trials as the “found solution” can be adopted. In either case, however, if the target value is unknown, it inevitably takes a long time for the simulation compared with the case where the target value is known.

We investigated optimization of systematic traffic signal control using a GA. In this problem too, no target value of the optimal solution is given, and the best solution obtained during a large number of trials is treated as the “optimal



solution” [1]. Because the optimal solution is not knowable, it is considered appropriate to describe the solution obtained in these cases as a quasi-optimal solution in the strict sense; however, here we call it the optimal solution without making such a distinction.

Optimal control of road traffic signals, which are one form of social infrastructure, is important from the standpoint of safeguarding the environment and achieving energy savings, and there has been a great deal of research on this topic [2]. There are three control parameters of traffic signals: cycle length, split, and offset. Cycle length is the time required for a signal to display a complete sequence, that is, the time from start of the current green indication to the start of the next green indication, generally expressed in seconds. Split is the length of time assigned to the green indication during the cycle time, expressed in seconds or as a percentage of the cycle length. Offset represents the time lag for commencing the green indication between adjacent signals, expressed in seconds or as a percentage of the cycle length. Cycle length and split are control parameters of independent signals, but in systematic control involving adjacent signals and multiple signals, offset is used as an additional control parameter. In systematic control optimization, there are also issues such as division of the road system and assigning system-wide cycle lengths for individual roads; however, the basic concepts have already been shown [3] and are now in practical use.

From the combination of offsets at each intersection, the problem of obtaining the offset combination that ensures smooth traffic flow is the key issue. However, the process of determining the signal control parameters for oversaturated traffic conditions is influenced by government policies on traffic management, and therefore, the subject of this study was non-saturated traffic.

The model road studied here has 21 intersections. If it is assumed that the offset can be set to 100 values from 0% to 99%, the solution search space becomes enormous; therefore, the optimal offset combination was sought using a GA. In the search for this optimal solution, the solution was calculated with the average travel time in both traffic flow directions defined as an evaluation metric, and its reciprocal defined as the fitness value. However, as described above, because a target value of the fitness, which is a function for evaluating the solution, is unknown, in order to obtain a solution in a practical amount of time, a method for reducing the simulation time is required.

In this study, noting that the fitness distribution when offset patterns are randomly given follows a probability distribution, we decided to obtain the fitness target value using road data of the simulated road as parameters, and we successfully reduced the simulation time.

2 Searching for an optimal (quasi-optimal) offset using a genetic algorithm

The number of offset combinations becomes enormous as the number of signals increases. For example, for a road having 21 intersections, as shown in fig. 1, the



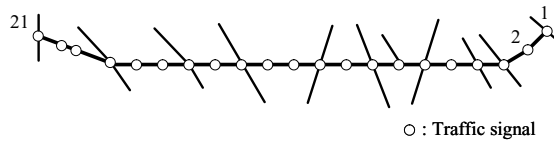


Figure 1: The model road.

offset for each intersection has 100 patterns, from 0% to 99%, and the number of combinations is 100^{21} , so it is impossible in practice to evaluate all of them to find the optimal offset combination. In this study, therefore, we investigated the use of a genetic algorithm (GA).

On the other hand, one simple method currently used in calculating the offsets in practice involves limiting offsets to basically two values, 0% and 50%, using the concept of link round-trip required travel time [3]. In brief, the blockage or passage of traffic flow according to the offset is obtained by simulation, and the flow is plotted on a time-space graph; the planner then decides the offset pattern by iterating a method involving fine adjustment of the offsets particularly for intersections where flow is obstructed, followed by re-simulation. However, this planning requires manpower, takes time, and the obtained offsets may or may not be optimal. Because the workload involved is high, once this offset pattern is set, unless there is a serious problem, it remains in use for a long time.

To overcome this problem, our goal is to realize optimal signal control according to the traffic flow by automatically calculating the offset pattern in response to changes in traffic flow conditions. In traffic control centers, measured values of the traffic flow volume etc. are obtained every 2.5 to 5 minutes. Even at mid-sized traffic control centers, there are 500 to 1000 signal intersections, and even unit roads for systematic control, known as sub-areas, may include as many as 100 to 200 roads, and in practice, the time required to search for optimal offsets is multiplied by the number of roads. In addition, central processing units in the traffic control center are not used exclusively for offset searching, and therefore, there is a demand to substantially reduce the time required to search for the optimal solution in order to allow them to function in real-time systems.

The simulator employed in this study is what we call "Smooth-Kun 2". The Architecture of the simulator is shown in fig. 2. It is constructed of a traffic simulator that calculates fitness values for vehicles traveling under the control of signals according to the supplied offsets, as well as a GA engine that performs a sequence of processes, such as evaluation, selection, crossing, and mutation based on the fitness values obtained from the traffic simulator.

As the fitness values calculated by the traffic simulator, we used the reciprocal of the average travel time (ATT), which is the average value of the travel time of all vehicles. In the offset evaluation, rather than just a simple travel time, a journey with a small number of stops in response to signals for the same travel time was considered to be better, and to take this into account in the calculation of the average travel time [3], 30 s was added to the travel time as a penalty for each

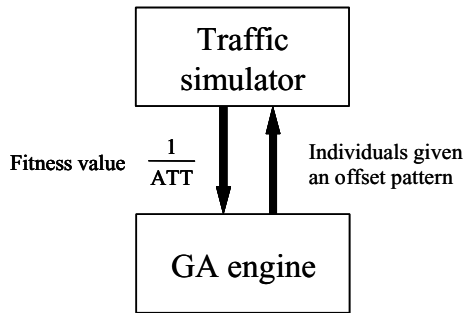


Figure 2: Architecture of simulator "Smooth-Kun 2".

stop. The fitness value is given by:

$$\text{fitnessvalue} = 1/ATT \quad (1)$$

where

$$ATT = \frac{1}{N} \sum_{i=1}^N (T_i + 30 \times S_i) \quad (2)$$

Here, the parameters are:

T_i : travel time of the i -th vehicle (including actual waiting time at stop signals),

S_i : number of stops of the i -th vehicle, and

N : total number of vehicles in both directions.

The genes in the GA corresponded to the arrangement of offsets for each signal, and the optimal offset combination was sought by iterating for a fixed number of generations. An example of the results is shown in fig. 3 [1].

From fig. 3, in the GA search, it was found that the ATT converges to an optimal solution as the number of generations increases, but the rate at which it converges is generally unclear. Ideally, all roads should converge to an optimal solution within the observation period of the traffic conditions; however, because of issues such as the computational performance of the computers that implement the GA, and because the target value of the evaluation function is essentially unknown, there is a limit to the reduction in search time, while necessarily depending on the method used, such as iteratively executing the calculation and terminating it when an improvement in the solution is no longer seen.

From the example results in fig. 3, the average travel time (ATT), which corresponds to the reciprocal of the fitness value, converges to about 610 s after about 550 generations. This means that, if a target value of about 610 s could be estimated prior to searching, there is a possibility of substantially reducing the search time. To implement this technique, it is necessary to estimate a value of the ATT (target ATT) which gives the target fitness value, indicating the degree of fitness of the solution, as the basis for terminating the GA calculation from the given conditions.

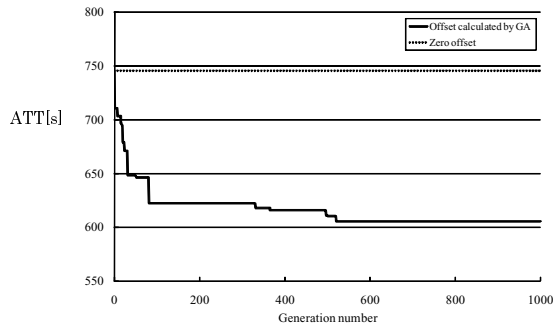


Figure 3: ATT versus generation number.

3 Possibility of estimating target ATT

To estimate the target ATT, it is necessary to ascertain the kind of distribution exhibited by the ATTs for random offset patterns in the solution space for the target road. If the details of the ATT distribution can be determined, there is a possibility of estimating in advance the optimal ATT obtained in the optimal offset search time as the target ATT based on that distribution. From this standpoint, we attempted to derive a general expression for estimating the target ATT, by considering an application to searching for an optimal offset. In estimating this ATT distribution, a generalized approach is preferred, rather than focusing on a specific road. We decided to vary the parameters characterizing the road (called traffic parameters) to derive an estimation formula that does not depend on the road. The traffic parameters used in this study are explained here. There are many traffic parameters that are considered to have an influence on the ATT characteristics. Of these, we focus on the total road length, the number of signals, the average link length, the link length variation, and the traffic demand factor, which expresses the level of traffic flow. The cycle length and the travel speed can be dependently inferred from the influence of the average link length with the concept of link round-trip required travel time, and therefore, they were treated as fixed values. Also, regarding the split, because its proportional relation with the stopping probability is clear, it too was treated as a fixed value and was excluded from the scope of analysis here. As the link length variation in this study, the standard deviation of the link length was defined as the link deviation. ATT increases as the total road length L increases, which also influences the number of intersections. Therefore, to perform a comparison by varying the traffic parameters, it is not appropriate to make direct use of the ATT value. Thus, we decided to normalize the ATT to enable comparison between them. To normalize the ATT, the travel time required to pass through all intersections without encountering any stop signals was defined as 0, and the travel time required when a vehicle is stopped by the signals at all intersections was defined as 1.

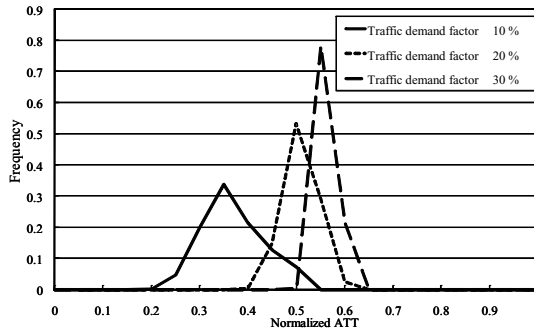


Figure 4: Normalized ATT distributions obtained by varying the traffic demand factor.

3.1 Effect of demand factor

The demand factor was varied in steps of 10% in the range 10-30%, and the ATT distribution for randomly produced offsets was normalized and is plotted in fig. 4. From fig. 4, we find that the distributions for all demand factors take a form resembling a normal distribution. The number of vehicle groups experiencing saturated flow tends to increase as the demand factor increases; therefore, because the signal stopping probability increases, the distribution also shifts in the direction of increasing normalized ATT.

3.2 Effect of link deviation

The ATT distribution was calculated while varying the link deviation. The ATT was calculated for roads in the case where the link deviation was set to 0 m and approximately 200 m, from a reference time. Because there are multiple combinations of link length when the link deviation is approximately 200 m, we examined the following two patterns, the results of which are shown in fig. 5:

Link pattern 1 (0:362:766:260:612m)

Link pattern 2 (0:317:790:311:582m)

From fig. 5, all the distributions approximated normal distributions, and when the link deviation was large, the variation was also large, and the standard deviation of the distribution increased.

3.3 Effect of average link length and number of intersections

Next, we looked at the ATT distribution for random offsets when the average link length was varied. Because the average link length is determined by two parameters, namely, the total road length and the number of intersections (number of links), both of them were varied to obtain the ATT distribution.

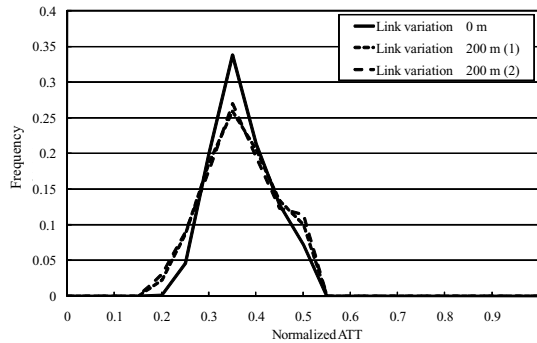


Figure 5: Normalized ATT distributions obtained by varying the link variation.

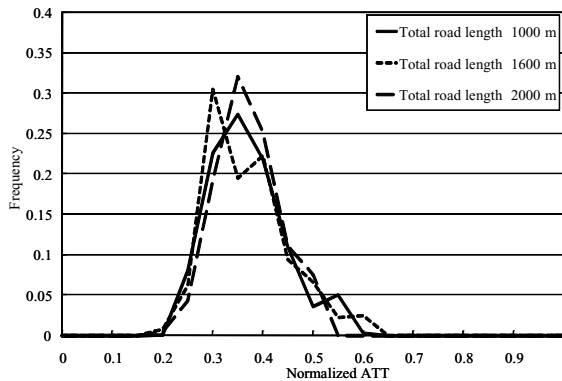


Figure 6: Normalized ATT distributions obtained by varying the total road length.

First, variation of the average link length was achieved by varying the total road length. The number of links was fixed at a reference value, and the total road lengths were set to 1400 m, 1600 m, and 1800 m. The resulting average link lengths were 350 m, 400 m, and 450 m. The ATT was calculated for the road in each case, and the results obtained for the normalized ATT are shown in fig. 6.

From fig. 6, the distributions took a form resembling a normal distribution also for the case of varying the total road length. Moreover, although the center value did not vary, the distributions showed a smaller variation as the total road length increased. In contrast, the results for the ATT distributions when the number of links was varied, keeping the road length unchanged, and when the average link length was varied are shown in fig. 7. In this analysis, by keeping the total road length fixed and setting the number of links to 5 and 8, the average link length was varied, taking values of 400 m and 250 m, respectively.

From fig. 7, the frequency distribution of the normalized ATT could be approximated by a normal distribution also in the case where the number of links was

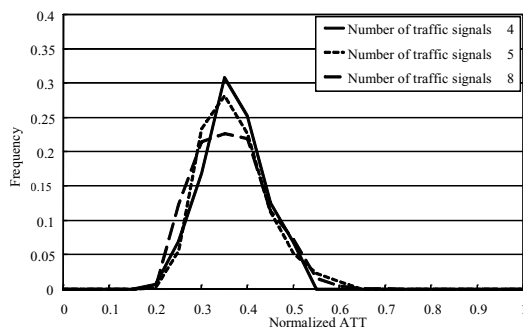


Figure 7: Normalized ATT distributions obtained by varying the number of links.

varied. In addition, although the distribution showed a large variation as the number of links was increased, no large variation was observed in the center value.

From the above results, for the ATT distributions in the case where random offsets were applied and four parameters were varied, the center value of the distribution changed when the demand factor was varied; however, when the demand factor was constant, even when the road length, the number of links, etc. were varied, the center value was substantially the same, and only the variation of the distribution was affected. These findings were elucidated by introducing the concept of normalized ATT.

We found that the ATT distribution obtained when giving random offset patterns was a normal distribution. To derive a target ATT on the basis of the normal distribution, the center value and standard deviation of the distribution should be obtained. Thus, we examined a method of calculating the center value and standard deviation from the demand factor and road parameters.

4 Normal distribution and parameters

Given that the normalized ATT can take a form resembling a normal distribution, the data is plotted on a normal probability plot to determine the center value and the standard deviation. The normal probability plot, in which the normalized ATT is plotted on the horizontal axis and the cumulative probability (Z value) on the vertical axis, shows that the straighter the line, the closer the data fits a normal distribution. The intersection with the horizontal axis represents the mean of the distribution, and the gradient corresponds to the standard deviation of the distribution.

First, fig. 8 shows the results of the normal probability plot for the ATT distribution obtained while varying the demand factor.

From fig. 8, the ATT distributions obtained while varying the demand factor shows substantially linear characteristics in the region ± 2 . The standard deviation increases as the demand factor is increased, and the mean also increases. Additionally, as the demand factor increases, the distribution becomes more linear.

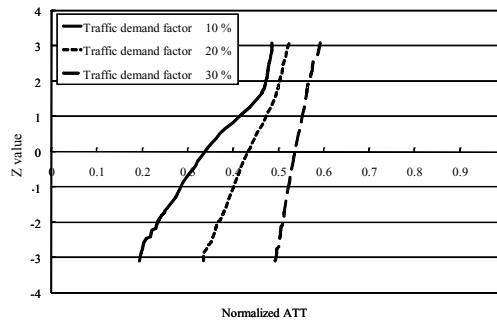


Figure 8: Z value characteristics obtained by varying the traffic demand factor.

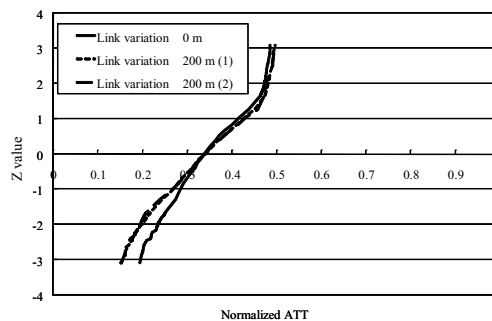


Figure 9: Z value characteristics obtained by varying the link variation.

Similarly, the results of the normal probability plot for the ATT distribution obtained while varying the link deviation are shown in fig. 9.

From fig. 9, there is a small variation in the distribution when the link deviation is 0, and the mean remains at approximately the same value even when the deviation differs. For a link deviation of 200, an interesting result was obtained in that, although the deviation is the same, two cases occur where the value differs, but they are both the same distribution. In addition, we performed analysis by changing the number of links without changing the ATT distribution or the road length, while varying the average link length according to the total road length, and determined the relationships between them. Based on the results of these analyses, we attempted to derive a general formula for estimating the ATT distribution.

5 Derivation of formula for estimating ATT distribution by multivariate analysis

From the results of our investigation thus far, the ATT distribution obtained by varying the traffic parameters that we focused on can be considered a normal

distribution. Therefore, to derive a formula for estimating the ATT distribution, two formulae should be derived, one for estimating the mean and one for estimating the standard deviation. Assuming that the various traffic parameters are independent of each other, we attempt to estimate the mean and standard deviation of the ATT distribution using multivariate analysis in which each estimated value is simply linearly combined.

Defining X_1 as the average link length, X_2 as the number of links, X_3 as the link deviation, and X_4 as the demand factor, the formula derived for the mean, μ , obtained by multivariate analysis is shown in (3), and the formula derived for the standard deviation, σ , is shown in (4):

$$\mu = 7.49 \times 10^{-5} X_1 - 0.01 X_2 - 2.37 \times 10^{-6} X_3 + 0.76 X_4 + 0.36 \quad (3)$$

$$\sigma = -0.2 \times 10^{-3} X_1 - 0.59 \times 10^{-2} X_2 - 4.22 \times 10^{-5} X_3 - 0.15 X_4 + 0.21 \quad (4)$$

In order to determine the validity of these formulae, we investigated the estimated values obtained with the derived formulae and the distribution obtained with the traffic simulator using the coefficient of determination R^2 . The results of investigation for the mean of the ATT distribution are shown in fig. 10, and the results of investigation for the standard deviation are shown in fig. 11.

From these results, in the means of the ATT distribution, a high correlation of about $R^2 = 0.97$ was obtained between the calculated value and the estimated value. In addition, a high correlation of about $R^2 = 0.91$ was also obtained for the standard deviations.

From these findings, the assumption that the various traffic parameters are all independent is valid, demonstrating that the mean and standard deviation of the normalized ATT can be calculated using the traffic parameters.

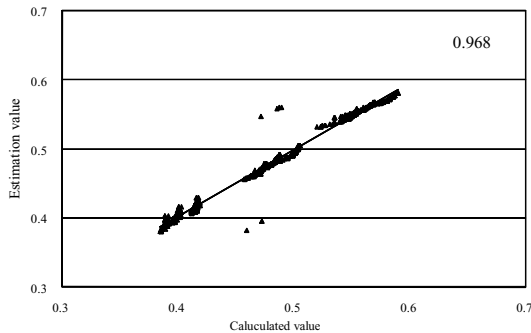


Figure 10: Comparison of the mean values.

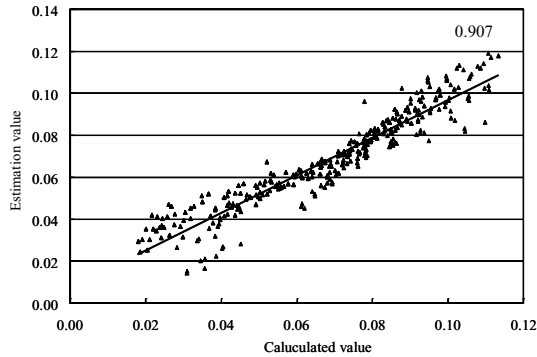


Figure 11: Comparison of the standard deviation value.

6 Verification of method for reducing search time by estimating ATT distribution

We verified the validity of applying this method to a target road containing 21 intersections. Because the converged value of the solution in the conventional search method is unknown, we adopted a method in which the calculation is terminated at a fixed number of generations, and the smallest solution obtained by performing this operation multiple times is selected. The total number of generations using the GA search was 10,000.

In contrast, when estimating based on conditions involving the road parameters and the vehicle demand factor, the estimated optimal solution -3 away from the mean was 0.224 in the normalized ATT. In terms of the ATT, this corresponds to 610 s, and therefore, we performed a search with 610 s set as the termination condition for the GA search. As a result, the search time of about 85 minutes required in the conventional method was reduced to about 4.5 minutes. It was thus possible to achieve a substantial reduction in the search time, demonstrating the validity of the proposed method.

7 Conclusions

The aim of our study was to substantially reduce the calculation time required to search for optimal control conditions of traffic signals using a genetic algorithm (GA). Our strategy was to derive a target ATT serving as a basis for terminating the GA calculation, from given conditions for the road under study. We examined the tendency of the ATT distributions in cases where four traffic parameters were varied for random offsets. As a result, for all of the parameters investigated, the ATT distributions were found to be close to normal distributions. Then, based on this finding, we attempted to derive estimation formulae for the mean and standard deviation of the ATT distribution by multivariate analysis, assuming that

the ATT distribution is a normal distribution. As a result, although the demand factor needed to be considered in categories, it was possible to achieve a high correlation between the calculated value and the estimated value for the mean and the standard deviation. Also, as a result of conducting a case study based on these results, a substantial reduction in search time compared with a conventionally used GA search method was achieved, demonstrating the high potential of the proposed method. In future it should be possible to improve the precision by conducting more detailed case studies. In addition, once the precision is validated in practice, we will examine the possibility of applying our method to real-time processing software in actual systems for systematic control of traffic signals.

References

- [1] S. Takahashi, H. Kazama, T. Fujikura and H. Nakamura, "Adaptive Search of an Optimal Offset for the Fluctuation of Traffic Flow Using Genetic Algorithm", *The transactions of the Institute of Electrical Engineers of Japan. D.*, vol. 123, no. 3, 2003, pp. 204–210.
- [2] G.E. Shoup and D. Bullock, "Dynamic Offset Turning Procedure Using Travel Time Data", *Transportation Research Record*, vol. 1683, no. 99-1086, 1999, pp. 84–94.
- [3] *Manual on Traffic Signal Control (Revised Edition)*, Japan Society of Traffic Engineers, 2006.

