Signal Timing Optimization for Corridors with Multiple Highway-Rail Grade Crossings Using Genetic Algorithm

Yifeng Chen
AECOM, Southfield, MI, yfchen@huskers.unl.edu

Laurence Rilett
University of Nebraska - Lincoln, lrilett2@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/civilengfacpub

Part of the Transportation Engineering Commons

http://digitalcommons.unl.edu/civilengfacpub/140

This Article is brought to you for free and open access by the Civil Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Civil Engineering Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Research Article

Signal Timing Optimization for Corridors with Multiple Highway-Rail Grade Crossings Using Genetic Algorithm

Yifeng Chen\textsuperscript{1} and Laurence R. Rilett\textsuperscript{2}

\textsuperscript{1}AECOM, Southfield, MI 48034, USA
\textsuperscript{2}Nebraska Transportation Center, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

Correspondence should be addressed to Yifeng Chen; yfchen@huskers.unl.edu

Received 26 August 2017; Revised 22 January 2018; Accepted 22 February 2018; Published 16 April 2018

Academic Editor: Taha H. Rashidi

Copyright © 2018 Yifeng Chen and Laurence R. Rilett. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Safety and efficiency are two critical issues at highway-rail grade crossings (HRGCs) and their nearby intersections. Standard traffic signal optimization programs are not designed to work on roadway networks that contain multiple HRGCs, because their underlying assumption is that the roadway traffic is in a steady-state. During a train event, steady-state conditions do not occur. This is particularly true for corridors that experience high train traffic (e.g., over 2 trains per hour). In this situation, the non-steady-state conditions predominate. This paper develops a simulation-based methodology for optimizing traffic signal timing plan on corridors of this kind. The primary goal is to maximize safety, and the secondary goal is to minimize delay. A Genetic Algorithm (GA) was used as the optimization approach in the proposed methodology. A new transition preemption strategy for dual tracks (TPS\_DT) and a train arrival prediction model were integrated in the proposed methodology. An urban road network with multiple HRGCs in Lincoln, NE, was used as the study network. The microsimulation model VISSIM was used for evaluation purposes and was calibrated to local traffic conditions. A sensitivity analysis with different train traffic scenarios was conducted. It was concluded that the methodology can significantly improve both the safety and efficiency of traffic corridors with HRGCs.

1. Introduction

In the US, there are an estimated 209,655 highway-rail grade crossings (HRGCs) [1]. Because the roadway and railway intersect at-grade, these crossings may have negative impacts on both safety and operations of the adjoining roadway network. Approximately 90 to 95\% of railroad-related fatal injuries occur at these facilities [2, 3]. In addition, many US roadways, particularly in the Midwest and Western states, were built adjacent to the railways. Consequently, roadway intersections near highway-rail grade crossings (IHRGCs) need to be operated in such a way that queues that are on or close to the HRGCs may be cleared prior to a train arrival. These special traffic signal operating instructions at an IHRGC are referred to in the literature as train preemption [3].

In addition, the presence of a train can have a negative effect on traffic at an IHRGC. Those movements that are not allowed can result in long queues, which might, in turn, disrupt other movements. This problem is compounded for roadway corridors that parallel a railway line, as there may be multiple IHRGCs that are impacted during the same train event. In many parts of the US, train traffic is so high that steady-state conditions never develop on the roadway network. For example, some cities in Nebraska have over 150 train events per day. In these situations, the frequent signal preemptions at IHRGCs can negatively impact both signal coordination and traffic flow, with the result that the capacity of the corridor is severely reduced [4].

There is a need for a new traffic signal optimization methodology for corridors with multiple IHRGCs that have a large amount of train traffic. This paper proposes a simulation-based optimization methodology for traffic signal coordination and preemption operations along corridors with multiple IHRGCs, which seeks to maximize safety and minimize delay. The optimization methodology is based on a Genetic
Algorithm (GA) approach, and a new transition preemption strategy for dual tracks (TPS_DT) is integrated into the proposed methodology.

2. Background

2.1. Signal Optimization. Signal optimization algorithms are techniques that systematically generate signal timing plans, evaluate fitness or objective functions (e.g., delay, bandwidth efficiency, and throughput) by using a simulation or analytic model, and find the best signal timing plan based on a predetermined criterion [5]. TRANSYT-7 F [6] and SYNCRHO [7] are two commonly used programs for optimizing signal timings, and both use macroscopic-deterministic models and a delay-based objective function. In contrast, the PASS programs (e.g., PASSER I through PASSER V) are a series of bandwidth-based programs that apply macroscopic models and use a variety of optimization algorithms including exclusive search, interference minimization, and Genetic Algorithms [5].

During the last decade, stochastic optimization has become increasingly popular in the traffic signal timing field. This new approach commonly uses Genetic Algorithms (GA) [8] and state-of-the-art microscopic simulation tools, such as CORSIM [9] and VISSIM [10], to optimize signal timings. Genetic Algorithms (GA) can be classified as global-optimization algorithms and are based on the natural selection principle and evolution process [8]. The procedure starts with a randomly generated set of chromosomes, each of which represents a potential solution to the problem under consideration. For signal timing optimization, the solutions are a combination of the four signal timing parameters: cycle length, phase sequences, offset, and green-time splits. The process of evolution is an iterative process that consists of four main steps: evaluation, selection, crossover, and mutation. First, an objective function is used to evaluate the performance of individual chromosomes. Then, the fittest individual chromosomes are selected as parents to generate offspring individuals. Mutation and crossover are variation-inducing operators that create offspring individuals from parents and preserve diversity within the population. The evolution process continues through many generations until certain termination criteria are reached. The criteria for the termination of the GA procedures can be either a convergence threshold of the fitness value, that is, the difference between the best fitness value and the average fitness value of the current population, or the maximum number of generations to evolve.

Many researchers have applied GA in signal timing optimization. Foy et al. [11] first used GA for signal timing design. They used a GA to optimize phase sequence and phase splits for an arterial of four intersections by minimizing total delay. It was found that GA can find near optimal signal timings. Park et al. [12, 13] were the first to use GA to simultaneously optimize all four signal timing parameters (cycle length, offsets, splits, and phase sequence) of an arterial with four closed spaced intersections. The results showed that GA generated better timing plans than TRANSYT-7 F for low and high demand scenarios and derived signal timing plans that were comparable to those from TRANSYT-7 F for the medium demand scenario. GA has also been effectively applied to solving more advanced problems of signal optimization, such as transit signal priority (TSP) [14–17]. These studies have proved that GA can improve the performance of transit vehicle, while minimizing the negative impacts of TSP on vehicular traffic.

Although there are some studies on optimizing signal timing plan with signal preemption at IHRGCs [4, 18], few literature has been found to use GA in this matter.

2.2. Preemption Strategies at HRGCs. The standard practice of signal preemption at HRGCs is known as the standard preemption (SP) strategy [19]. It provides vehicles and pedestrians with minimum warning time. To provide the minimum warning time, the algorithm may abruptly truncate phases that are active, including the pedestrian phases, which may be problematic if pedestrians are in the crosswalk. To improve the safety, many studies have been conducted on preemption operation of traffic signals at HRGC with active warning devices [20–26]. The transition preemption strategy (TPS) [20] developed by the Texas Transportation Institute (TTI) uses an additional detector upstream of the SP detector to provide vehicles and pedestrians with advance warning preemption time (AWPT). This greatly reduces the number of pedestrian phases that are truncated, but the prediction errors in train arrival times can still result in pedestrian phase truncations or high vehicle delay [21]. The improved transition preemption strategy (ITPS) was then developed based on TPS, and a new train arrival prediction algorithm was incorporated in this strategy [21–23]. The prediction algorithm can increase the accuracy of train arrival prediction by frequently updating the train arrival time and estimating the bounds of prediction error. Compared to the TPS, the ITPS assigns higher importance and more green time to the phases that will be blocked during the preemption before the preemption sequence is initiated, rather than those that are served during the preemption. Simulation studies [22, 23] have illustrated that ITPS can significantly improve both the safety and efficiency of HRGC operations. The primary limitation of the ITPS is that it was designed for only one HRGC and for a single train event. Corridors with multiple HRGCs, dual-track, and simultaneous train event were not considered. A new transition preemption strategy for dual tracks (TPS_DT), which was partially based on the ITPS logic, was developed by Chen [27]. This strategy is appropriate for corridors with multiple HRGCs and that have dual rail tracks, which may have multiple trains passing in both directions. It has been shown that the TPS_DT can improve both safety and efficiency of isolated intersections near HRGCs [27].

3. Methodology

The proposed optimization methodology consists of two components: an optimization module and a simulation module. Figure 1 illustrates the architecture of the proposed methodology.

The optimization module optimizes signal timings, subject to the constraints of the signal controller logic and
preemption logic. Note that any optimization program may be utilized. In this paper, a Genetic Algorithm (GA) [8] was used for the optimization due to its capability of solving complex, nonconvex optimization problems such as signal timing optimization.

The simulation module consists of a microsimulation model of the roadway/railway network, a train arrival prediction system, and a preempted logic algorithm for HRGCs. VISSIM [10] was used to model the roadway/railway network, because of its capability of modeling multimodal traffic, for example, vehicles, pedestrians, and trains. The performance evaluator in VISSIM outputs traffic system performance metrics including average delay, number of stops, and queue length as well as train information including track speed, train length, and train travel time.

A train arrival time prediction model is integrated in the simulation model to predict train arrival times at downstream HRGCs. Note that any appropriate prediction model, from the basic kinematic model used in the constant warning time (CWT) system [19] to more advanced statistical models [28], may be applied. In this paper, the basic kinematic model based on distance and instantaneous train speed [19, 28] was used for train arrival time estimation.

A traffic signal preemption logic algorithm is also integrated in the simulation model and controls the preemption operation of the HRGCs in the simulation model. In this paper, both the standard preemption logic and the transition preemption strategy logic for dual tracks (TPS_DT) are applied for comparison purposes. Detailed information regarding the TPS_DT may be found elsewhere [27].

The methodology is iterative in nature. First, the GA optimizer generates candidate signal timing plan parameter sets and sends them to the simulation model for performance evaluation. Second, the simulation model, in conjunction with the train arrival prediction and traffic signal preemption logic, is run using the signal timing plan parameter sets obtained from the optimization model. Third, the performance of each candidate signal timing plan set is evaluated during the simulation and sent back to the GA optimizer. The parameter sets with top performance values are identified among these candidate plan sets and used to generate the candidate plan sets for the next iteration. Lastly, the iteration process stops when certain stopping criteria (e.g., maximum iteration numbers in this paper) are met. The best signal timing plan is identified by the end of the iterations.

The proposed signal timing optimization methodology was implemented in a combined program of MATLAB and Visual Basic (VB). The MATLAB provides toolboxes to perform the GA routines, including the generation of the initial population and the selection of parent chromosomes, crossover, and mutation. A VB program was developed to enable the data exchange between the GA program and the VISSIM simulation model.

### 3.1. Optimization Objectives

#### 3.1.1. Safety-Related Objective

The primary objective of the methodology is to improve the safety of pedestrians in the corridor. In this paper, the number/percentage of pedestrian phase truncations is used as the MOE to evaluate the safety performance of the methodology. Equation (1) is used to calculate the probability of a pedestrian phase being active during a given cycle. Based on this equation, the probability of a pedestrian phase being truncated at the start of the track clearance phase is 99.9% for the current SP system, if the pedestrian volume is equal to 400 ped/h. In the simulation, the pedestrian volumes for all pedestrian phases were set to 400 ped/h to ensure that a pedestrian phase is active and truncated by the SP system. It is hypothesized that the transition preemption strategy, TPS_DT [27], integrated in the signal controller logic at HRGCs, will significantly decrease or eliminate the pedestrian phase truncations at the intersections near the HRGCs during preemption

\[
P_r(n \geq 1) = 1 - P_r(n = 0) = 1 - e^{-\lambda t},
\]

where

- \(n\) is number of pedestrians arriving during the affected time period of every cycle
- \(P_r(n \geq 1)\) is probability of one or more pedestrians arriving during the affected time period of every cycle
- \(\lambda\) is pedestrian arrival rate (ped/h)
- \(t\) is length of time period that a pedestrian can affect the next pedestrian phase (=115 s)
- \(e\) is natural base of logarithms (=2.71828···)

#### 3.1.2. Efficiency-Related Objective

The secondary goal of the methodology is to improve the efficiency of the study corridor. In this paper, vehicular delay is chosen as the MOE for evaluating the efficiency of the signal timing plan. Equation (2) is used to calculate the average delay of a given intersection \(j\), where the delay on a given movement

\[
\text{Delay}_j = \frac{\sum_{i=1}^{m} t_{ij} \cdot d_{ij}}{\sum_{i=1}^{m} d_{ij}}
\]

where

- \(m\) is number of movement in an intersection
- \(t_{ij}\) is travel time on the movement \(i\) in the direction of movement \(j\)
- \(d_{ij}\) is demand on the movement \(i\) in the direction of movement \(j\)

**Figure 1:** Architecture of the proposed methodology.
where

- $C$ is cycle length (s)
- $M$ is total number of intersections in the corridor
- $G_{ij}$ is green time for phase $i$ at intersection $j$
- $I_{ij}$ is interchange time for phase $i$ at intersection $j$, including amber and all-red clearance time
- $\theta_j$ is offset of the signal at intersection $j$ (s)
- $\min G_{ij}$ is minimum green time for movement $i$ at intersection $j$ (s)
- $\max G_{ij}$ is maximum green time for movement $i$ at intersection $j$ (s)
- $C_{\min}$ is minimum cycle length (s)
- $C_{\max}$ is maximum cycle length (s)

Equations (5) and (6) indicate the barrier constraints of NEMA signal structure; (7) is the cycle length constraint, which ensures that the sums of green phase times and interchange times (yellow change plus all-red clearance time) equal the cycle length. Equation (8) indicates the minimum and maximum values of the phase green time for each intersection. Equation (9) confines the upper and lower bounds of the cycle length. Equation (10) indicates that the offset should lie between zero second and the cycle length. Lastly, (11) limits the green time phase, cycle length, and offset to nonnegative integer values.

### 3.2. Decision Variables for Optimization

The goal of the signal optimization with preemption is to improve the safety and efficiency of the signal timing operation of a highway-railroad corridor before, during, and after preemption events. Therefore, the decision variables for the optimization include both basic signal timing parameters and preemption-related signal timing parameters.

There are four basic signal timing parameters for normal signal operation without preemption: cycle length, phase split, phase sequence, and offset. All four basic parameters are optimized simultaneously using the GA [12, 30]. For the optimization of signal timings during the preemption process, the phase sequence is usually kept fixed, because changing the phase sequence during the preemption will make the process more complicated and consequently increase the confusion of drivers. Therefore, cycle length, phase split, and offset are chosen as the optimization parameters in GA. For coordinated-actuated signal control, the phase split is calculated as the maximum green time plus the yellow change and all-red intervals [31]. Because the yellow and all-red intervals are usually constant parameters, the maximum green time is optimized. The force-off points and permissive periods for coordination operation are then calculated accordingly after the maximum green times of all phases are determined [31]. Because the yellow and all-red intervals are constant, the maximum green time is the decision variable that is optimized. Once the maximum green time for each phase is identified, the force-off points and permissive periods for coordination operation are determined [31].
For the target IHRGCs, several preemption-related parameters are involved in the optimization process. These parameters are identified from the preemption stage at HRGCs, which can be divided into the four stages [18] as follows:

1. Advance warning stage begins upon detection of a train by the advanced detector, and when the transition preemption strategy, like TPS_DT, is activated. In this stage, the advance preemption warning time (APWT) for east bound \( (r_{EB}) \) and west bound \( (r_{WB}) \) are the decision variables.

2. Track clearance phase stage lasts from the time the SP starts until the train arrives at the crossing. In this stage, the track clearance time (TCT) is the decision variable that is optimized.

3. Preemption dwell stage starts from gate closure (gate down) to gate opening (gate up). No parameter is optimized in this stage, as the gate movement is controlled by the railroad.

4. Restoring stage lasts from gate opening until returning to normal operation. The exit phase duration (EP) is the decision variable that is optimized. Table 1 summarizes the two groups of decision variables for the signal optimization in the paper.

### Table 1: Decision variables for signal optimization.

<table>
<thead>
<tr>
<th>Group 1: basic signal parameters</th>
<th>Group 2: preemption-related parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Cycle length ((C))</td>
<td>(i) APWT ((r_{EB})) and ((r_{WB}))</td>
</tr>
<tr>
<td>(ii) Maximum green ((G))</td>
<td>(ii) Track clearance time ((TCT))</td>
</tr>
<tr>
<td>(iii) Offset ((O))</td>
<td>(iii) Exit phase duration ((EP))</td>
</tr>
</tbody>
</table>

\(1\) Applies to all signalized intersections in corridor. \(^*\) Applies to all signalized IHRGCs in corridor.

3.3. Generation of Decision Variables in GA. In this paper, a fraction-based decoding scheme [12] was applied in the GA to generate candidate basic signal timing parameter sets. This method uses fractional variables to assign available green times to phases on a prorated basis and is designed to accommodate the traffic signal control constraints. The following sections describe the decoding scheme.

#### 3.3.1. Generation of Basic Signal Timing Parameter Sets. For an eight-phase intersection, five to seven fractional variables are needed to calculate three basic signal timing parameters: cycle length, green split, and offset, depending on whether there are overlaps between the phases in the two rings. In the case of a phase with overlaps, seven fractional variables, denoted as \( f_1, f_2, f_3, f_4, f_5, f_6, f_7 \), are needed. Figure 2 illustrates the decoding theme for an eight-phase intersection with overlaps between the two phasing rings of the NEMA dual-ring structure. Among these variables, \( f_1 \) is used to calculate cycle length, \( f_2 \) splits the cycle length into mainstreet and cross-street times, and \( f_3 \) through \( f_6 \) determine the splits of the eight phases in the NEMA dual-ring structure.

The fraction to determine the offset of the signal to the master clock is \( f_7 \).

In the GA, the fractional variables for an intersection are displayed as a vector, for example, \( \{f_1; f_2; f_3; f_4; f_5; f_6; f_7\} \). Each fractional variable is encoded as an agent, and all agents constitute the chromosome of a signal timing solution [27, 32]. The agents are then decoded to signal timing parameters using (12) to (16).

Equation (12) is used to calculate the cycle length. The minimum and maximum cycle lengths in the equation are chosen based on the range of cycle lengths from the time-of-day (TOD) signal timings for the studied corridor.

$$ C_j = C_j^{\text{min}} + \text{Int} \left[ \left( C_j^{\text{max}} - C_j^{\text{min}} \right) * f_1 \right] $$

\( \forall j = 1, \ldots, M, \)

where

- \( C_j^{\text{min}} \) is minimum cycle length (s), 90 s for the study corridor
- \( C_j^{\text{max}} \) is maximum cycle length (s), 120 s for the study corridor
- \( \text{Int} \) is a rounding function to obtain integer cycle length.

The phase times of the main and cross-street phases are calculated as follows:

$$ P_j^{\text{main}} = \text{maximum} \left( \text{mp}_j^1 + \text{mp}_j^2, \text{mp}_j^3 + \text{mp}_j^4 \right) $$

$$ + \text{Int} \left[ \left( C_j - MP_j \right) * f_2 \right], $$

$$ P_j^{\text{cross}} = \text{maximum} \left( \text{mp}_j^5 + \text{mp}_j^6, \text{mp}_j^7 + \text{mp}_j^8 \right) $$

$$ + \text{Int} \left[ \left( C_j - MP_j \right) * (1 - f_2) \right], $$

$$ \text{MP}_j = \text{maximum} \left( \text{mp}_j^1 + \text{mp}_j^2 + \text{mp}_j^3 + \text{mp}_j^4, \text{mp}_j^5 + \text{mp}_j^6 + \text{mp}_j^7 + \text{mp}_j^8 \right), $$

where

- \( P_j^{\text{main}} \) is sum of phase times of the main-street phases for intersection \( j \)
- \( P_j^{\text{cross}} \) is sum of phase times of the cross-street phases for intersection \( j \)
- \( \text{mp}_j^i \) is minimum splits (green time plus yellow and all-red intervals) of phase \( i \) at intersection \( j \), \( i = 1, 2, \ldots, 8 \)
- \( \text{MP}_j \) is sum of minimum phase split time for intersection \( j \).
By applying the fraction variables $f_3$ through $f_6$, the phase times (splits) of phase 1 through phase 8 are calculated using the following equations:

$$P_{ij} = mp^i_j + \text{Int}\left\{\left[(C_j - MP_j) \times f_2\right] \times f_3\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_2 + \text{Int}\left\{\left[(C_j - MP_j) \times f_2\right] \times (1 - f_3)\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_3 + \text{Int}\left\{\left[(C_j - MP_j) \times (1 - f_2)\right] \times f_4\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_4 + \text{Int}\left\{\left[(C_j - MP_j) \times (1 - f_2)\right] \times (1 - f_4)\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_5 + \text{Int}\left\{\left[(C_j - MP_j) \times f_2\right] \times f_5\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_6 + \text{Int}\left\{\left[(C_j - MP_j) \times f_2\right] \times (1 - f_5)\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_7 + \text{Int}\left\{\left[(C_j - MP_j) \times (1 - f_2)\right] \times f_6\right\}$$

$$\forall j = 1, \ldots, M,$$

$$P_{ij} = mp^i_8 + \text{Int}\left\{\left[(C_j - MP_j) \times (1 - f_2)\right] \times (1 - f_6)\right\}$$

$$\forall j = 1, \ldots, M, \quad (14)$$

where

$$P_{ij}$$ is phase time of phase $i$ for intersection $j$.

After obtaining the splits of the 8 phases, the green time of the phases is calculated by subtracting the amber and all-red intervals from the split of each phase, as shown in

$$G_{ij} = P_{ij} - a_{ij} - r_{ij} \quad \forall i = 1, 2, \ldots, N, \ j = 1, 2, \ldots, M, \quad (15)$$

where

$G_{ij}$ is green time for phase $i$ of intersection $j$.

The fractional variable $f_7$ is used to generate the offset $O_j$ of the signal at intersection $j$. As shown in (16), $O_j$ is calculated as a fraction of the cycle length

$$O_j = \text{Int}\left\{(C_j \times f_2)\right\} \quad \forall j = 1, \ldots, M, \quad (16)$$

where

$O_j$ is offset for intersection $j$.
Table 2: Preemption-related optimization parameters for the target IHRGCs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Name in the GA program</th>
<th>Min. (s)</th>
<th>Max. (s)</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Track clearance time at intersection 3</td>
<td>TrackClrTime33</td>
<td>15</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>Track clearance time at intersection 4</td>
<td>TrackClrTime35</td>
<td>15</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>Track clearance time at intersection 5</td>
<td>TrackClrTime44</td>
<td>15</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>Advanced preemption time for EB at intersection 3</td>
<td>APT33_EB</td>
<td>35</td>
<td>130</td>
<td>6</td>
</tr>
<tr>
<td>Advanced preemption time for WB at intersection 3</td>
<td>APT33_WB</td>
<td>35</td>
<td>130</td>
<td>6</td>
</tr>
<tr>
<td>Advanced preemption time for EB at intersection 4</td>
<td>APT35_EB</td>
<td>35</td>
<td>130</td>
<td>6</td>
</tr>
<tr>
<td>Advanced preemption time for WB at intersection 4</td>
<td>APT35_WB</td>
<td>35</td>
<td>130</td>
<td>6</td>
</tr>
<tr>
<td>Advanced preemption time for EB at intersection 5</td>
<td>APT44_EB</td>
<td>35</td>
<td>220</td>
<td>7</td>
</tr>
<tr>
<td>Advanced preemption time for WB at intersection 5</td>
<td>APT44_WB</td>
<td>35</td>
<td>65</td>
<td>5</td>
</tr>
<tr>
<td>Exist phase duration at intersection 3</td>
<td>ExitPhase33</td>
<td>10</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>Exist phase duration at intersection 4</td>
<td>ExitPhase35</td>
<td>10</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>Exist phase duration at intersection 5</td>
<td>ExitPhase44</td>
<td>10</td>
<td>35</td>
<td>5</td>
</tr>
</tbody>
</table>

practices at HRGC in the United States [33]. The maximum values of advanced preemption warning times (APWT) are based on the distance of the advance detectors from the intersections and the fastest train speed in the field. The minimum APWT should not be less than the constant warning time plus phase interchange time (amber plus all-red clearance time). In this paper, the constant warning time is set to 25 s for all three HRGCs, and the phase interchange time is set to 6 s for the phases of all three intersections. This adds up to 31 s. To be conservative, a minimum APWT value of 35 s is chosen. For the duration of the exit phase, the minimum value is the current exit phase duration of 10 s, while the maximum value is the maximum green time for the same phase (Phase 4) of normal signal operation. Table 2 lists the preemption-related signal timing parameters for each intersection, their minimum and maximum values, and the number of genes for each agent representing them in GA.

4. Experiment Design

4.1. Study Network and Optimization Corridor. The study network for this paper is a 2.4 km by 3.2 km urban road network in Lincoln, Nebraska. The network is bounded on the west by North 27th Street, on the east by North 48th Street, on the north by Superior Street, and on the south by Holdrege Street. A map of the study network is illustrated in Figure 3. A 3.2 km section of the BNSF railroad goes through the northeast and southwest corners of the network. This is a dual-track mainline railroad, which crosses North 27th Street with an overpass, and North 48th Street with an underpass, respectively. The overpass at the 27th Street and the underpass at the 48th Street form the geographic boundaries of the network. Within the boundaries, there are three HRGCs, located at North 33rd Street, Adams Street, and North 44 Street. They are marked as “H1,” “H2,” and “H3” in Figure 3, respectively. In addition, there are three signalized intersections near the three HRGCs: 33rd Street and Cornhusker Highway, 35th Street and Cornhusker Highway, and 44th Street and Cornhusker Highway. They are marked as intersections 3, 4, and 5 in Figure 3 and referred to as the target intersections in Figure 3 and throughout the paper.

In Figure 3, the study corridor for signal optimization is indicated by the red solid line with arrows, which includes the Cornhusker Highway (i.e., Nebraska State Highway 6), the BNSF railroad line, and the three HRGCs along the Cornhusker Hwy. The Cornhusker Highway runs parallel to the BNSF railroad east of the Adams Street HRGC. In addition to the three target intersections, the optimization corridor also includes another three signalized intersections on the corridor: 27th St and Cornhusker Highway (Intersection 1), 29th St and Cornhusker Highway (Intersection 2), and 48th St and Cornhusker Highway (Intersection 6).

The optimization corridor was chosen because (1) Cornhusker Hwy is a major arterial in Lincoln, an alternative route to I-80 between Lincoln and Omaha, Nebraska, and the traffic volume, especially truck traffic, is high; (2) there are approximately 50 to 70 trains per day passing on the BNSF railroad line and through the HRGCs, and the train volume is increasing [27, 34]; (3) the Adams Street HRGC has been rated as one of the most hazardous HRGCs in Lincoln by the FRAs Web Accident Prediction System (WBAPS) [34]; and (4) this corridor serves as the UNL HRGC test bed system and is instrumented with train detectors [34]. Therefore, safety and delay are major concerns along this corridor.
4.2. Simulation Model Setup. To create the simulation network in VISSIM, photographs of the corridor obtained from Google Maps were imported into the program. The highway lanes and railroad tracks were coded based on the mapped images. The characteristics of the network, including traffic volumes, speed limits, detector lengths and locations, lane width, and track width, were obtained from the Google map or provided by the Public Works Department of the city of Lincoln. The most recent signal timing settings were also obtained from the Public Works Department of Lincoln and coded in Vehicle Actuated Programming (VAP), which is an add-on module of VISSIM for signal control [10]. For the train preemption strategy, both SP and TPS_DT logic algorithms were coded as a module of the VAP logic and emulated in the simulation. The simulation network also includes pedestrian crossings and pedestrian signals at the three target IHRGCs. Morning peak hour traffic volumes from the traffic count data in 2009 and common driving behavior parameters in VISSIM, such as minimum headway and maximum deceleration rate, were used to calibrate the model. The Geoffrey E. Havers (GEH) statistic [35, 36] was used as the objective function for the calibration. The GEH statistic for the calibrated model was 5.19, which indicates that the calibration is acceptable, considering the large scale of the study network. The model was validated using the turning ratio data in 2009 at five of the study network’s major intersections. The validation accuracy was 83%, in terms of mean average percentage error (MAPE). In the model calibration process, train length and train speed data collected from the field were incorporated in the model and used as the train length and speed distribution for the simulation. During the simulation, train speed and length were randomly assigned to a train from the train length and speed distributions. More details regarding the model calibration and validation can be found elsewhere [27].

4.3. Simulation Scenarios. Based on the train volume distribution collected from the field, three train volumes, that is, 1 train/h, 3 train/h, and 5 train/h, were used, representing low, high, and extremely high train volumes, respectively [27]. In the study corridor, trains can arrive from the east, west, or simultaneously from both directions. In the case of trains traveling in both directions, there are chances that two trains from different directions may occupy one HRGC simultaneously, depending on their speed and length.

A total of nine simulation scenarios, that is, three approach scenarios by three volumes, were used in the optimization. Table 3(b) lists the nine optimization scenarios, where the GA algorithm along with the TPS_DT [27] was applied. Table 3(a) lists the corresponding nine baseline scenarios where the current signal timings and the SP logic were used. Each scenario is labeled in the form of “x-y-z,” where “x” represents optimization/baseline scenarios (e.g., “1” represents optimization scenarios, and “0” is baseline scenarios), “y” represents the train direction (e.g., “E” represents EB train, “W” represents WB train, and “B” represents trains from both directions), and “z” represents the number of the trains in each direction (e.g., 1, 3, or 5 trains/h).
4.4. Simulation Duration. The simulation duration was set to 3600 seconds to simulate the morning peak hour traffic. In the preliminary simulation study, the network was found to reach a steady-state condition approximately 600 s after the simulation started [27]. Based on this, the analysis period of the simulation was set to 2700 s for all the simulation scenarios, starting at 600 s and ending at 3300 s of the simulation period. In each scenario, the train departure times were fixed so that a comparison between baseline and optimization scenarios could be performed under the same traffic conditions. The other advantage of the fixed train departure time is that the first train can be guaranteed to enter the network after the warm-up time of 900 s. For the scenarios with 1 train/h, the train departs at 1800 s, which is in the middle of the simulation period. For the scenarios with 3 trains/h and 5 trains/h, the first train departs at 600 s and the last train departs at 3000 s. The intervals between two consecutive trains are 1200 s and 600 s for 3 trains/h and 5 trains/h, respectively. The departure times of the trains are listed in Table 4.

4.5. GA Parameter Settings. In this paper, a MATLAB toolbox for GA [32] was used to implement the GA routines. A population size of 30 individual chromosomes per generation and a maximum number of 30 generations were applied to the GA optimizer coded in MATLAB. First, a binary matrix of 30 by 108 was generated in MATLAB as the initial population, using the “crtbp” function in the MATLAB toolbox. The corresponding functions for the GA routines in the toolbox are “select” for selection/reproduction of chromosomes, “recombine” for crossover, and “mut” for mutation. The uniform method [8] is used in the function “recombine” for the crossover routine. In this function, the crossover rate was set to 0.7. The mutation rate was set to 0.01 in the function of “mut”.

5. Simulation Results and Findings

5.1. Safety Metric Evaluation. Tables 5–7 show the number of pedestrian phase cutoffs and the percentages of cutoffs in the preemption events for the nine optimization scenarios for EB, WB, and both directions, and their corresponding baseline scenarios. Reductions in the percentages of pedestrian phase cutoffs in the optimization scenarios ranged from 59% to 100%, compared to their baseline scenarios. In the scenarios with one train (i.e., scenario 1-E-1 and scenario 1-W-1), the pedestrian truncation has been eliminated, because the pedestrian truncation has been eliminated, because the TPS_DT module was initiated for every preemption event. For the other scenarios, it was observed that more than one train passed the HRGCs during the simulation, and the TPS_DT module could not be initiated for all the preemption events due to the limitation of the VAP module in VISSIM [27]. This resulted in pedestrian phase cutoffs, because only the SP was initiated in those events. With more advanced controller capabilities, it is hypothesized that the pedestrian phase cutoffs in all the scenarios would be eliminated.

5.2. Delay Metric Evaluation. After the optimization process was complete, an evaluation of the optimized signal timing plans was conducted. To control the stochastic variability in VISSIM simulation, 50 multiple random-seeded simulation runs were performed with the optimized signal timings from the nine optimization scenarios. Another 50 simulation runs with the same set of random seeds were performed for the respective nine baseline scenarios with the current signal timing plan. This ensures that traffic pattern and train characteristics were identical for each optimization scenario and its corresponding baseline scenario. In this section, the evaluation results are shown and discussed.

The MOEs for delay evaluation at three levels, the average delay of the three target intersections, the average delay of the study corridor, and the average delay of the study network were output from the simulation. Mean values of the MOEs from 50 random-seeded simulation runs were then calculated, and a one-tail paired-t test was applied to compare the MOEs of the optimization scenarios with those of the respective baseline scenarios. Table 8 compares the average MOE values for the optimization scenarios and the respective baseline scenarios, by calculating the differences in percentage between them. The p values between the optimization scenarios and the respective baseline scenarios were also calculated for the paired t-test, to identify a
significant difference in delay between the optimized and baseline scenarios at the 0.05 significance level. The null hypothesis ($H_0$) of the paired $t$-test was that the average delay in the baseline scenarios is equal to that in the optimization scenarios, while the alternative hypothesis was that the average delay in the baseline scenarios is significantly greater than that in the optimization scenarios.

It can be seen in Table 8 that all optimized signal solutions result in a decrease of delay on the three target intersections and the whole corridor. There are statistically significant improvements on average delay of the three target intersections at the 5% significance level in eight of the nine scenarios. The only exception is scenario 1-W-3, where there was an improvement, but this improvement was not statistically significant at the 5% significance level. On average, there is a 14.3% reduction in the average delay of the target intersections. At the corridor level (see Table 8(b)), the optimized signal timing plans resulted in significant decreases of the average delay in all nine scenarios at a 95% confidence level. Over the nine scenarios, there is a 10.2% improvement in the average corridor delay. The improvement of delay for the three target intersections near HRGCs is higher than that for the whole corridor. This is probably because the signal timing settings of the other three intersections in the corridor (27th St and Cornhusker Hwy, 29th St and Cornhusker Hwy, and 48th St and Cornhusker Hwy) were not included in the objective function, and this averages out the improvement at the three target intersections.

On the other hand, the average network delay values of the optimization scenarios are significantly higher than those of the respective baseline scenarios. As shown in Table 8(c), there is an average 5% increase in the network delay. This indicates that there is a tradeoff between improving corridor performance and network performance. However, it can be argued that the safety and efficiency issues of the corridor traffic are more important than the efficiency of the whole network during the preemption, and thus more weight should be placed on the corridor with HRGCs.

### 6. Concluding Remarks

In this paper, an optimization methodology and a GA-based program were developed for corridors with multiple HRGCs. Different simulation scenarios were performed in order to evaluate the methodology. Nine simulation scenarios with different combinations of train volumes and directions were designed for morning peak hour simulation. The optimized signal timings were evaluated with 50 multiple simulation runs and then compared with those of the baseline scenarios with current signal timings.

The GA-based optimization program combined with the new transition preemption strategy for dual tracks (TPS_DT) was found to significantly improve both safety and efficiency of the corridor. Regarding the safety issue of pedestrians at the HRGCs, the pedestrian phase cutoffs of the TPS_DT preemption strategy were reduced by 60% to 100%. It is hypothesized that pedestrian cutoffs could be eliminated if TPS_DT were properly implemented in the field. Meanwhile, the improvement of efficiency is evaluated with the MOEs of the vehicular delay at three different levels: (1) the average delay of the target intersections has been reduced by approximately 14%; (2) there is an approximately 10% reduction in the average delay at the corridor level including all six intersections; and (3) the average delay of the whole study network increased by approximately 5% because the objective function of the optimization was focused on improving the efficiency of the corridor traffic, instead of the traffic in the whole network. The results indicate that there is a tradeoff
between the network efficiency and the corridor efficiency for the signal optimization.

It is recommended that the proposed GA-based program be applied to other corridors with multiple HRGCs, to test whether the approach is geographically transferrable. In this paper, all simulations were performed using AM peak hour traffic volumes, and fixed train arrival times and train volumes were used. For future research, different levels of traffic demand, including the non-peak hour traffic demand, and empirical train arrival times and volumes can be used to see if changes in these assumptions affect the results in terms of safety and efficiency. Since the timetable for trains in North America has extreme variability and may not be accessible to the researchers, it would be more realistic to collect train data from the field and simulate the train arrivals based on empirical data. In this case, a 24-hour or longer simulation would be needed.

**Conflicts of Interest**

The authors declare that there are no conflicts of interest regarding the publication of this paper.
Table 8: Comparison of multiple run results between optimization and baseline scenarios.

(a) Average delay of 3 intersections near HRGCs.

<table>
<thead>
<tr>
<th>Number of trains</th>
<th>Simulation scenarios</th>
<th>Baseline scenarios</th>
<th>Optimization scenarios</th>
<th>( ((2) - (1))/(1) )</th>
<th>( P(T \leq t) )</th>
<th>( H_0^*; ) ( (1) = (2) )</th>
<th>( H_1^*; ) ( (1) &gt; (2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 train in EB</td>
<td>Scenario 0-E-1 versus scenario 1-E-1</td>
<td>60.27</td>
<td>48.30</td>
<td>-19.9%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>3 trains in EB</td>
<td>Scenario 0-E-3 versus scenario 1-E-3</td>
<td>66.93</td>
<td>61.82</td>
<td>-7.6%</td>
<td>0.01</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>5 trains in EB</td>
<td>Scenario 0-E-5 versus scenario 1-E-5</td>
<td>89.96</td>
<td>76.48</td>
<td>-15.0%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>1 train in WB</td>
<td>Scenario 0-W-1 versus scenario 1-W-1</td>
<td>62.73</td>
<td>51.86</td>
<td>-17.3%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>3 trains in WB</td>
<td>Scenario 0-W-3 versus scenario 1-W-3</td>
<td>71.10</td>
<td>68.46</td>
<td>-3.7%</td>
<td>0.07</td>
<td>Accept ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>5 trains in WB</td>
<td>Scenario 0-W-5 versus scenario 1-W-5</td>
<td>96.91</td>
<td>83.50</td>
<td>-13.8%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>1 train in EB &amp; WB</td>
<td>Scenario 0-B-1 versus scenario 1-B-1</td>
<td>68.28</td>
<td>57.27</td>
<td>-16.1%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>3 trains in EB &amp; WB</td>
<td>Scenario 0-B-3 versus scenario 1-B-3</td>
<td>69.30</td>
<td>57.23</td>
<td>-17.4%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>5 trains in EB &amp; WB</td>
<td>Scenario 0-B-5 versus scenario 1-B-5</td>
<td>129.67</td>
<td>108.29</td>
<td>-16.5%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>79.46</td>
<td>68.14</td>
<td>-14.3%</td>
<td>*</td>
<td>Reject ( H_0 ) at the 5% significance level.</td>
<td></td>
</tr>
</tbody>
</table>

(b) Average corridor delay.

<table>
<thead>
<tr>
<th>Number of trains</th>
<th>Simulation scenarios</th>
<th>Baseline scenarios</th>
<th>Optimization scenarios</th>
<th>( ((2) - (1))/(1) )</th>
<th>( P(T \leq t) )</th>
<th>( H_0^*; ) ( (1) = (2) )</th>
<th>( H_1^*; ) ( (1) &gt; (2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 train in EB</td>
<td>Scenario 0-E-1 versus scenario 1-E-1</td>
<td>72.57</td>
<td>65.12</td>
<td>-10.3%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>3 trains in EB</td>
<td>Scenario 0-E-3 versus scenario 1-E-3</td>
<td>77.27</td>
<td>73.60</td>
<td>-4.8%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>5 trains in EB</td>
<td>Scenario 0-E-5 versus scenario 1-E-5</td>
<td>88.41</td>
<td>85.15</td>
<td>-3.7%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>1 train in WB</td>
<td>Scenario 0-W-1 versus scenario 1-W-1</td>
<td>74.77</td>
<td>69.90</td>
<td>-6.5%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>3 trains in WB</td>
<td>Scenario 0-W-3 versus scenario 1-W-3</td>
<td>77.95</td>
<td>75.52</td>
<td>-3.1%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>5 trains in WB</td>
<td>Scenario 0-W-5 versus scenario 1-W-5</td>
<td>91.46</td>
<td>84.59</td>
<td>-7.5%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>1 train in EB &amp; WB</td>
<td>Scenario 0-B-1 versus scenario 1-B-1</td>
<td>77.27</td>
<td>66.30</td>
<td>-14.2%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>3 trains in EB &amp; WB</td>
<td>Scenario 0-B-3 versus scenario 1-B-3</td>
<td>86.12</td>
<td>70.56</td>
<td>-18.1%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>5 trains in EB &amp; WB</td>
<td>Scenario 0-B-5 versus scenario 1-B-5</td>
<td>108.05</td>
<td>86.42</td>
<td>-20.0%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>83.76</td>
<td>75.24</td>
<td>-10.2%</td>
<td>*</td>
<td>Reject ( H_0 ) at the 5% significance level.</td>
<td></td>
</tr>
</tbody>
</table>

(c) Average network delay.

<table>
<thead>
<tr>
<th>Number of trains</th>
<th>Simulation scenarios</th>
<th>Baseline scenarios</th>
<th>Optimization scenarios</th>
<th>( ((2) - (1))/(1) )</th>
<th>( P(T \leq t) )</th>
<th>( H_0^*; ) ( (1) = (2) )</th>
<th>( H_1^*; ) ( (1) &gt; (2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 train in EB</td>
<td>Scenario 0-E-1 versus scenario 1-E-1</td>
<td>325.51</td>
<td>356.70</td>
<td>9.6%</td>
<td>0.00</td>
<td>Reject ( H_0 )</td>
<td></td>
</tr>
</tbody>
</table>
Average delay (s/veh) | Baseline scenarios | Optimization scenarios | Difference (%) | P(T <= t) | H_0* | H_1 |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 trains in EB</td>
<td>Scenario 0-E-3 versus scenario 1-E-3</td>
<td>346.20</td>
<td>375.99</td>
<td>8.6%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>5 trains in EB</td>
<td>Scenario 0-E-5 versus scenario 1-E-5</td>
<td>384.98</td>
<td>414.63</td>
<td>7.7%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>1 train in WB</td>
<td>Scenario 0-W-1 versus scenario 1-W-1</td>
<td>330.45</td>
<td>364.08</td>
<td>10.2%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>3 trains in WB</td>
<td>Scenario 0-W-3 versus scenario 1-W-3</td>
<td>352.17</td>
<td>377.88</td>
<td>7.3%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>5 trains in WB</td>
<td>Scenario 0-W-5 versus scenario 1-W-5</td>
<td>403.77</td>
<td>429.83</td>
<td>6.5%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>1 train in EB &amp; WB</td>
<td>Scenario 0-B-1 versus scenario 1-B-1</td>
<td>335.90</td>
<td>358.91</td>
<td>6.8%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>3 trains in EB &amp; WB</td>
<td>Scenario 0-B-3 versus scenario 1-B-3</td>
<td>376.58</td>
<td>369.28</td>
<td>-1.9%</td>
<td>0.03</td>
<td>Reject H_0</td>
</tr>
<tr>
<td>5 trains in EB &amp; WB</td>
<td>Scenario 0-B-5 versus scenario 1-B-5</td>
<td>472.22</td>
<td>450.10</td>
<td>-4.7%</td>
<td>0.00</td>
<td>Reject H_0</td>
</tr>
</tbody>
</table>

Average 369.75 388.60 5.1% | 5.1% | 0.03 | Reject H_0 |

*Reject H_0 at the 5% significance level.

References


