UMEC
URBAN MOBILITY & EQUITY CENTER

Final Report

Optimized Development of Urban Transportation Networks

Paul Schonfeld

Maryland Transportation Institute

1173 Glenn Martin Hall

University of Maryland, College Park

MD 20742

Date

May 11, 2019

MORGAN STATE UNIVERSITY
GROWING THE FUTURE · LEADING THE WORLD

UNIVERSITY OF MARYLAND
18 56

VirginiaTech
Invent the Future

ACKNOWLEDGMENT

Disclaimer

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Optimized Development of Urban Transportation Networks		5. Report Date May 11. 2019
		6. Performing Organization Code
7. Author(s) *Include ORCID #* Dr. Paul Schonfeld orcid.org/0000-0001-9621-2355		8. Performing Organization Report No.
9. Performing Organization Name and Address Maryland Transportation Institute 1173 Glenn Martin Hall, University of Maryland College Park, MD 20742		10. Work Unit No.
		11. Contract or Grant No. 69A43551747123
12. Sponsoring Agency Name and Address US Department of Transportation Office of the Secretary-Research UTC Program, RDT-30 1200 New Jersey Ave., SE Washington, DC 20590		13. Type of Report and Period Covered Final
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This report presents improved methods for planning and scheduling interrelated improvements in transportation networks. Due to the complexity of the relevant evaluation functions, which cannot be optimized with classical calculus techniques, the proposed methods rely on customized genetic algorithms for optimizing the selection, sequencing and scheduling of the interrelated alternatives. Three applications to urban transportation networks are presented in journal papers which are included in appendices. The papers demonstrate the applicability of the proposed methods to urban road networks, to intersections in urban road networks and to the development of urban rail transit networks.

17. Key Words : Network Development, Network Optimization, Project Scheduling, Interrelated Projects		18. Distribution Statement	
19. Security Classif. (of this report) : Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 68	22. Price

Optimized Development of Urban Transportation Networks

Paul Schonfeld
University of Maryland, College Park

Final Report for the University Mobility and Equity Center

May 11, 2019

Table of Contents

Abstract

Executive Summary

Appendix 1 – Selecting and Scheduling Interrelated Projects

Appendix 2 – Selecting and Scheduling Link and Intersection Improvements in Urban Networks

Appendix 3 – Optimal Development of Rail Transit Networks

Abstract

This report presents improved methods for planning and scheduling interrelated improvements in transportation networks. Due to the complexity of the relevant evaluation functions, which cannot be optimized with classical calculus techniques, the proposed methods rely on customized genetic algorithms for optimizing the selection, sequencing and scheduling of the interrelated alternatives. Three applications to urban transportation networks are presented in journal papers which are included in appendices. The papers demonstrate the applicability of the proposed methods to urban road networks, to intersections in urban road networks and to the development of urban rail transit networks.

Executive Summary

In this project methods were developed for planning, evaluating and scheduling improvements in transportation networks in order to optimize the development of such networks in response to evolving demand and societal objectives. The work was performed at the University of Maryland, College Park, in the years 2017 and 2018, with funding from the University Center for Mobility and Equity led by Morgan State University, as well as from other sources. The work was directed by Professor Paul Schonfeld from the University of Maryland's Department of Civil and Environmental Engineering. Important contributors included his students Elham Shayanfar, Uros Jovanovic and Ya-Ting Peng, and Professor Zi-Chun Li from the Huazhong University of Science and Technology in Wuhan, China.

The problems of selecting and scheduling improvements in transportation networks are greatly complicated by the pervasive interrelations among candidate alternatives. In engineering economics and other fields the alternatives are classified as **(a) mutually exclusive, (b) independent, and (c) interrelated**. The alternatives are considered mutually exclusive if only one alternative may be chosen, the others being necessarily rejected. They are independent if the benefits **and** costs of each alternative do not depend on which other alternatives are selected or when the other alternatives are implemented. If the benefits **or** costs of alternatives depend on which others are selected and when all are implemented, the alternatives are classified as interrelated. While generally accepted methods for analyzing mutually exclusive and independent alternatives can be found in standard textbooks no such **general** methods are found for analyzing interrelated alternatives. Furthermore, even the methods that have been designed for analyzing interrelated alternatives in some **specific applications** have been deficient in their abilities to deal with complex interrelations, dissimilar types of alternatives, multiple uncertainties, scheduling decisions, realistic problem sizes and other important factors.

The deficiencies of methods for analyzing interrelated alternatives constitute a major gap in the state of the art in engineering economics, operations research, and related fields. This is especially unfortunate since interrelated alternatives pervade the world. For example, in transportation systems, which are the primary focus of our proposed study, improvements to a network's various links and nodes are interrelated partly because such improvements redistribute flows in networks. Each improved link may divert traffic from parallel links, shift congestion and capacity bottlenecks to other links in-series, reduce the need for other

improvements, and thus affect the benefits obtainable from improving other network components. Hence the benefits of various improvements may add up non-linearly. Some improvement projects may be synergistic while others may be largely wasted or even counterproductive (e.g., according to the Braess Paradox) when combined with other improvements.

Beyond interrelations due to non-linearly additive benefits (including some externalities), alternatives may be interrelated through their costs (e.g., through economies of jointly constructing several projects), their budget constraints and other financial relations, their constructability or operability requirements, political or equity considerations, and in other ways.

In addition, decisions regarding infrastructure maintenance or development are subject to substantial **uncertainties** regarding future demand or usage, costs, finances, implementation schedules, and future component performance (including capacity, delay, deterioration, and failures). Methods have been developed for dealing with uncertainties in capacity expansion and maintenance for infrastructure projects but these are far from adequate in dealing with realistic numbers of interrelated projects and their applicability is limited.

The three appendices of this report present three papers on the analysis of interrelated alternatives for transportation networks. The first paper (in Appendix 1) by E. Shayanfar and P. Schonfeld, entitled "Selecting and Scheduling Interrelated Projects: Application in Urban Road Network Investment," presents a metaheuristic method based on a genetic algorithm for optimizing network development problem. The metaheuristic approach is needed because for realistic problem sizes the objective function is very unsmooth and not solvable with either classical methods of mathematical analysis or with mathematical programming approaches. The paper shows how a genetic algorithm can be formulated and applied to efficiently solve this problem. In effect, the method consists of expressing all possible sequences for implementing alternatives as genetic chromosomes, translating the sequences into exact development schedules (in continuous time rather than discrete periods) by applying the binding constraints (which, in this case, are the budget constraints) and using a relatively simple traffic assignment algorithm to estimate traffic speeds and volumes throughout a multi-year analysis period for any development schedule. The traffic speeds and volumes can then be used to estimate other effectiveness measures, including travel times and user costs, throughout the analysis period.

Since heuristic methods do not guarantee that a global optimum is always found, the paper shows how a statistical test can be used to confirm that the infinitesimal probability of finding significantly better solutions than those found by the proposed heuristic method. Thus, it can be demonstrated that any errors due to the proposed algorithm are negligible compared to unavoidable errors in estimating inputs regarding the actual transportation system and its future demand characteristics.

The second paper (in Appendix 2) by U. Jovanovic, E. Shayanfar and P. Schonfeld, titled "Selecting and Scheduling Link and Intersection Improvements in Urban Networks" shows how the analysis, selection and scheduling of interrelated components in urban road networks

can be extended to include improvements at intersections, i.e., widening the intersections with additional lanes through them. To accomplish this, the traffic assignment model had to be adapted to analyze intersection flows and delays. This was accomplished by introducing into the previously used Frank Wolfe assignment algorithm pseudo-links for each turning and through movement at each intersection, e.g., 12 pseudo-links at each full four-leg intersection. Delays on the pseudo-links were estimated with a model developed by Akcelik.

The third paper (in Appendix 3) by Y. T. Peng, Z.C. Li and P. Schonfeld, titled "Optimal Development of Rail Transit Networks over Multiple Periods," shows how the analysis, selection and scheduling of interrelated network components can be extended to optimize the phased development of a rail transit network, In this problem it is assumed that the locations of rail lines and stations in the network are pre-determined. The remaining decisions are about which links and stations should be added at what time, depending mainly on demand growth, available external budgets and usable fare revenues from network segments that are already operating.

The methods developed and tested in this project are already usable for evaluating selecting and scheduling interrelated network improvement projects. Beyond the accomplishments of this project, desirable improvements would include improved consideration of uncertainties (e.g. in demand, costs, budgets and construction times) and extensions to multi-modal transportation systems.

Selecting and Scheduling Interrelated Projects: Application in Urban Road Network Investment

Elham Shayanfar and Paul Schonfeld

ABSTRACT

Decisions about the selection of projects, alternatives, investments, operating policies and their implementation schedules are major subjects in various fields including operations research, financial analysis, business management, engineering economy and transportation planning. In these various disciplines sufficiently good methods have been developed for planning and prioritizing projects when interrelations among those projects are negligible. However, methods for analyzing interrelated alternatives are still inadequate. We propose a combinatorial method for evaluating and scheduling interrelated road network projects. Specifically, this paper demonstrates how a traffic assignment model can be combined effectively with a Genetic Algorithm (GA) in a multi-period analysis to select and schedule road network projects while capturing interactions among those projects. The goal is to determine which projects should be selected and when they should be funded in order to minimize the present value of total system cost over a planning horizon, subject to budget flow constraints.

1. INTRODUCTION

Evaluating transportation infrastructure projects and determining which should be implemented at what time has been the subject of ongoing studies for decades. Commonly used evaluation practices aggregate linearly the project impacts in the objective function, which is then optimized. Such practices are often inadequate, especially for projects in transportation networks, since they disregard possible interrelations among projects due to non-linearly additive benefits, costs, budget constraints, constructability or operability requirements, and other possible factors. This paper deals with road expansion projects as an example of interrelated projects. However, the method proposed here for project selection and scheduling may be used to analyze interrelated alternatives in general cases if methods for evaluating objective functions are available.

In various disciplines sufficiently good methods have been developed for dealing with projects which are not interrelated. In general, alternatives are classified as (a) mutually exclusive, (b) independent and (c) interdependent or "interrelated". The alternatives are considered mutually exclusive whenever implementing one project automatically excludes the others. Alternatives are independent if their benefits and costs do not depend on which other alternatives are selected or when the other alternatives are implemented. Otherwise, the alternatives are classified as interdependent. Although generally accepted methods for analyzing mutually exclusive and independent alternatives are available in the literature, no such general methods are found for analyzing interrelated alternatives. Even the methods that have been designed for analyzing interrelated alternatives in some specific applications have been incapable of dealing with enough interrelations and realistic problem features.

The problem of evaluating and selecting interdependent alternatives exists in various fields including economics, operations research, business, management, transportation and portfolio management. In portfolio management, interrelations between choices (stocks) were identified and modelled as early as the 1950s in pioneering work by Markowitz (1952). Since then more recent studies have addressed the problem of portfolio selection among interdependent projects (Cruz et al., 2014; Li et al., 2016). However, the literature review shows both insufficient studies on this problem and lack of comprehensive applicable methods for real world problems especially in the field of transportation.

This study demonstrates how a relatively simple method, namely a traffic assignment algorithm, can be efficiently used to evaluate the objective function of an investment planning optimization problem and thereby compute the relevant interrelations among many projects that are implemented at various times. However, more complex methods for evaluating the objective functions, such as microscopic simulations, can also be combined with the Genetic Algorithm (GA) used here for optimizing the project selection and schedule. In recent years, meta-heuristics have been widely used for finding optimal or near-optimal solutions. The work presented in this paper is an extension of a previous study conducted by Shayanfar et al. (2016). That study applied three meta-heuristic algorithms including a GA, Simulated Annealing (SA) and, Tabu Search (TS) in seeking efficient and consistent solutions to the selection and scheduling problem. Its main contribution was to compare three meta-heuristics for this problem in terms of solution quality, computation time and consistency. The comparative analysis was especially useful in determining which algorithm was preferable in various circumstances. In summary, the results indicated that the GA yielded a better (lower total cost) solution than the other two algorithms and yielded the most consistent solutions (i.e. with the lowest coefficient of variation), indicating that different replications of the GA yield almost similar final solutions after sufficient iterations.

Therefore, the current paper incorporates the GA used in Shayanfar et al. (2016) while enhancing its assumptions and contributing to the literature in several ways. First, we demonstrate how a traffic assignment model can be combined effectively with a GA for

planning and prioritizing purposes while capturing more interactions among projects, i.e. beyond the previously considered pairwise interactions. Second, we modify the algorithms' assumptions to account for the possibility that candidate projects may become economically justified or unjustified after the implementation of previous projects. This may occur due to project interrelations and the possibility that the cost savings from completing a project are affected by earlier project implementations. Third, a multi-period analysis is incorporated in this study to distinguish between peak and off-peak traffic flows. Fourth, the budget constraint is reformulated to include possible internal funding from fuel taxes. Fifth, we assume that the demand changes over time during the planning horizon (growing exponentially in our example). Finally, we demonstrate this methodology on two example networks and present a statistical test of the goodness of the heuristic results. Generally, the methodology presented in this work should also be applicable to other prioritization problems with interrelated alternatives, which abound in transportation and other activities.

2. LTERATURE REVIEW

In engineering economics, a number of studies have developed methods to address the problem of project scheduling. Beenhakker and Narayanan (1975) formulated the scheduling problem as a simple integer program assuming projects are independent. The formulation maximized the total net benefit of all projects subject to a budget constraint. Chiu and Park (1998) proposed a capital budgeting model under uncertainty in which cash flow information was considered as a special type of fuzzy number. To prioritize fuzzy projects based on the present worth of each fuzzy project cash flow, a branch and bound procedure was suggested. Koc et al. (2009) proposed a model that forms an optimal priority list of projects, incorporating multiple scenarios for input parameters. For this purpose, a greedy heuristic algorithm was developed to create the prioritize list. Our research indicates that in the field of engineering economics and capital investment planning, the methods developed for selecting and scheduling do not adequately deal with possible interrelations among alternatives.

One of the first works we could find that considered interdependent alternatives was that of Markowitz (1952) on portfolio management. This study formulated a multi-objective function minimizing the sum of purchase cost and risks. In this case, a "dependence matrix" which captures two-way, three-way or n-way interrelations was introduced to model the interdependence among a set of choices. This method and its variants can also be found in more recent works. Dickinson et al. (2001) developed a model to optimize a portfolio of development improvement projects for the Boing Company. The authors used a dependence matrix to quantify the interdependencies among projects. Then a non-linear, integer program model was developed to optimize the project selection. Sandhu (2006) introduced a dependency structure matrix that captured the project logistic interdependencies. Durango-Cohen and Sarutipand (2007) formulated a quadratic programming for optimizing maintenance and repair (M&R) policies for transportation infrastructure systems. The quadratic objective of their work included the pairwise economic dependencies capturing the costs and benefits of improving adjacent facilities. Bhattacharyya et al. (2011) also considered n-way interdependencies in the Research and Development (R&D) project portfolio selection problem.

Two main issues arise from using a dependence matrix. First, as Disatnik and Benninga (2007) argue, the estimation and manipulation of a dependence matrix becomes computationally difficult as the project space grows. Second, the pairwise and n-way

dependencies do not completely represent the complex interrelations and fall short of the desired relations among alternatives. Instead of a dependence matrix, complete system models, such as queueing approximations (Jong and Schonfeld, 2001), equilibrium assignment (Tao and Schonfeld, 2005), microsimulation (Wang and Schonfeld, 2008) and neural networks (Bagloee and Tavana, 2012), are better suited for modeling interrelations. This section reviews the current literature on evaluating and prioritizing interdependent projects.

The SA algorithm developed by Bouleiman and Lecocq (2003) for the resource-constrained project scheduling problem aimed to minimize the total project duration. To this end, they replaced the conventional SA search scheme with a more novel design mindful of the specificity of the solution space of project scheduling problems. Tao and Schonfeld (2005) developed a GA to solve the Lagrangian problem, and optimized the selection of interdependent projects under cost uncertainty. They employed a traffic assignment model to evaluate the objective value of the Langragian problem and assess the project impacts. Similarly, Wang and Schonfeld (2005) developed a GA to solve the problem of selecting and scheduling interrelated lock improvements for a waterway network. They designed a microscopic waterway simulation model (i.e. which traced every vehicle movement) to assess the performance of the waterway system while evaluating the project interdependencies. Dueñas-Osorio et al. (2007) incorporated the interdependence response among network elements based on geographic proximity i.e. the response of one network given the state of another network was monitored for various levels of coupling among them. They studied the network response subject to external and internal disruptions such as attacks, lack of maintenance and breakdown due to aging. Their work indicated that responses that are destructive to networks are greater when interdependencies are considered after disruptions. Tao and Schonfeld (2007) developed island model variants of GA's for optimizing project selection and scheduling, and used these models to solve a stochastic optimization problem. Their work considered how uncertainties in travel times and construction costs affect total system costs.

Szimba and Rothengatter (2012) developed a framework for integrating the interdependence among infrastructure projects in classical benefit-cost analysis. They addressed the complexity of a large-scale interdependence problem by introducing a heuristic method to optimize the dynamic mixed integer program. In this approach, the number of projects and their interrelations were reduced stepwise, resulting in a fewer interdependence cases. They used two procedures to measure the magnitude of interdependencies. In the first, projects were added to a minimum network configuration. In the second, projects were deleted from a maximum network configuration. Bagloee and Tavana (2012) used the Traveling Salesman Problem (TSP) to formulate the prioritization problem. They used a Neural Network to consider the interdependence among projects, and developed a search engine influenced by Ant Colony (AC) hybridized with GA to optimize the problem. Li et al. (2013) developed a multi-commodity minimum cost network (MMCN) to evaluate the impact of projects, i.e. to estimate the benefits of projects through a life-cycle-cost analysis. They further proposed a hypergraph knapsack model to maximize these benefits for a set of interdependent projects. Rebiasz et al. (2014) developed a hybrid method which combined stochastic simulation with arithmetic on interactive fuzzy numbers and nonlinear programming. The goal was to solve the problem of capital budgeting, accounting for both stochastic and economic interdependency between projects.

Chen et al. (2015) reformulated the mixed network design problem (MNDP) to identify optimal capacity expansions of existing links and new link additions. Their model was designed to minimize the network cost in terms of the average travel time affected by the

expansion of existing links and the addition of new candidate links. In this case a surrogate-based optimization framework was proposed to solve the MNDP. Bagloee and Asadi (2015) developed a hybrid heuristic method to optimize the prioritization problem while considering demand uncertainties. They formulated the objective function as the reduction in users' travel time and, introduced a policy based on "gradient maximization" to find solutions. Tofighi and Naderi (2015) developed a mixed integer linear program to formulate the selection and scheduling of projects maximizing total expected benefits. They also proposed an ant colony algorithm to optimize the objective function. This paper defined the interdependencies among projects with a simple dependence matrix, which is insufficient in capturing the full interrelations among projects in transportation networks and various other complex systems.

3. PROBLEM FORMULATION

Roadway improvement projects are usually interrelated since delays at one link are affected by operations at other links, both upstream and downstream. Conceptually, if the capacity increases in one link of a network, congestion and average travel times tend to increase in other links that are "in series" with it and decrease in its "parallel" links. Therefore, the total cost saved from multiple projects is not a linear summation of savings from individual links. Additionally, the interrelation among links is reflected in our budget constraint since the budget is partly supplied by internal taxes, which may change after each project implementation, thus complicating this problem.

The objective function for problems such as prioritizing interrelated projects has a surface that is "noisy" (i.e. containing numerous local optima) and non-convex. Moreover, as the number of candidate projects increases, the problem's solution may soon exceed the capabilities of conventional mathematical optimization methods. Consequently, heuristic methods have become the preferred approach for solving such problems. In this study a GA is very useful in effectively finding near-optimal solutions for such a large solution space and noisy objective function. Our objective function is the net present value of total cost including both *(i)* total road user and *(ii)* total supplier cost subject to budget constraints. The goal is to specify which links should be selected for expansions in what order, and when they should be started and completed over the horizon period *T*.

Therefore, the formulated objective function minimizes the present value of total user and supplier cost, over a specified planning horizon, subject to a budget flow constraint over that entire horizon. In this context, the user cost is the total delay for users in the system multiplied by their value of time. The supplier cost is the present value of implementation costs for all projects. An additional improvement over some previous studies is the inclusion of project costs in the objective function. This is necessary since not all selected projects are guaranteed to fit in the budget and be implemented within the analysis period. In fact, some projects may be discarded from the sequence as they may become unjustified sometime during the analysis. Therefore, different solutions (i.e. different sequence of projects) may entail different project costs which should be considered in the objective function. The objective function Z to be minimized is the present value of total cost:

$$min\, Z = \sum_{j=1}^{T} \left\{ \frac{v}{(1+r)^j} \sum_{i=1}^{n_l} w_{ij} \right\} + \sum_{i=1}^{n_p} \frac{c_i x_i(t)}{(1+r)^t} \tag{1}$$

13

$$\begin{cases} x_i(t) = 0 & if \ t < t_i \\ x_i(t) = 1 & if \ t > t_i \end{cases}$$

In this formulation t_i is the time when project i is completed and ready for use while $x_i(t)$ is a binary variable specifying whether project i is finished by time t. In the objective function, w_{ij} denotes the travel time over link i in year j, and c_i is the present value of the cost of project i. n_p, n_l, v are the number of projects implemented, total number of links and value of time, respectively, while r is the interest rate.

In this problem an internal budget source is considered for funding future projects. Specifically, throughout the analysis period, fuel taxes collected from users are added to an external budget in determining the overall investment budget. This assumption is realistic, as fuel taxes and toll collections contribute substantially to highway improvement budgets. The internal budget is estimated as:

$$b(t_i)_{internal} = VMT(t_{i-1}) * f_r * f_c * f_t$$

where f_r, f_c, f_t denote fuel consumption rate (gal/veicle.mile), fuel cost (\$/gal), and gas tax rate (percentage of tax collected from dollar spent on gas) respectively. This formulation shows that fuel taxes collected in period t_{i-1} contribute to the budget available in period t_i. $VMT(t_{i-1})$ presents the vehicle miles travelled during the time project $i-1$ is completed. Jong and Schonfeld (2001) formulated the selection and sequencing problem by defining the decision variables as the completion time of projects. In this formulation the budget constraint is defined as follows:

$$\sum_{i=1}^{n_p} c_i x_i(t) \le \int_0^t b(t)dt, \quad 0 \le t \le T \tag{3}$$

More specifically, under a limited budget, which is continuously distributed over time, it is efficient to fund and complete projects one at a time, because the system gains immediate benefits as soon as each additional project is completed and ready for use. The budget constraint is almost invariably binding because, in actual cases, there are always some justifiable projects waiting for funding. In fact, funding multiple projects concurrently increases their completion time, meaning that their benefits are postponed. Therefore, considering budget limitations, it is preferable to avoid funding overlaps, and fully fund projects before starting to fund the next ones, and finish each project one at a time. It should be noted that construction times of projects may overlap even if their budget accumulation periods do not, if constrained budget flows can be shifted over time (e.g. through lending). Thus, the optimized schedule of each project is uniquely and easily determined from the optimized sequence by considering the budget flow.

To date, similar studies have assumed that the set of candidate projects remains unchanged throughout the analysis period, thus disregarding that due to interrelations, previous project implementations alter the benefits from completing succeeding projects, possibly making them economically unjustifiable. It is also possible that initially unjustifiable projects (i.e. with higher costs than benefits) may become economically desirable, e.g. after bottlenecks in networks are cleared. Accordingly, in this paper, the undesirable projects (i.e.

whose benefits < costs) are temporarily removed from the list of candidate projects, with the possibility of reentering the sequence after their benefits exceed their costs. In other words, the set of candidate projects is constantly updated, and acceptable projects may be replaced unacceptable ones at different stages of analysis.

4. EVALUATION MODEL

This paper applies the convex combination algorithm of Frank-Wolf (1956) as an evaluation model to assess the effects of each expansion project on the network. The Frank–Wolfe algorithm is an iterative first-order optimization algorithm for constrained convex optimization widely used for solving traffic assignment problems. In each iteration, the Frank–Wolfe algorithm considers a linear approximation of the objective function, and moves slightly towards a minimizer of this linear function. The algorithm starts with an initial flow x. Subsequently, each iteration performs a direction search by solving a linear approximation of the objective function which determines the step size and moves in that direction. Finally, the algorithm terminates when it satisfies a convergence criterion based on the similarity of successive solutions. In this case, the traffic assignment algorithm provides a relatively simple model for evaluating solutions (i.e. computing the objective function value), and estimating link travel times, speeds, volumes, and hence user costs.

5. OPTIMIZATION MODEL

In general, simulation methods are reserved for complex problems which are not solvable analytically. However, it may be computationally expensive to insert simulation modules directly into optimization loops. Hence, various approximation methods have been substituted for simulation (Dai and Schonfeld, 1998, Wei and Schonfeld, 1994). By now meta-heuristics, especially population-based ones such as GA's, along with faster computers, can solve complex optimization problems with unsmooth objective functions, even when simulation is used to evaluate the objective function (Balamurugan, 2006; Haq and Kennan, 2006; Wang and Schonfeld, 2005). In this paper a GA is used to find the optimal or near-optimal solution to the selection and scheduling problem. To test this approach, a Frank-Wolfe traffic assignment algorithm is used to compute the objective function. This algorithm can be replaced later with a detailed simulation model.

A GA (Genetic Algorithm) is a metaheuristic method that imitates the biological evolution and is based on the natural selection process (Michalewicz and Janikow, 1991). At first, GAs create a set of possible solutions which form the "initial population". This process mostly creates the initial population randomly. A string of encoded genes called a "chromosome" specifies each individual in the population. In this algorithm some individual solutions with the best "fitness" value (i.e. objective function value) are chosen to reproduce new offspring. This is usually a probabilistic process in which the individuals with better fitness values have a higher probability of being selected for creating the next generation. Then a series of mutation and crossover operators mate the selected solutions and change their attributes to maintain the population's diversity, and create the new generation (Golberg, 1989). In this study, each individual is defined as a string of numbers each corresponding to a specific project to be implemented (FIGURE 2). In addition to random order solutions, a greedy-order solution, a bottleneck-order solution form the initial population. In this context, the greedy-order solutions represent the sequence of projects ordered by their benefit-cost

ratio, disregarding their interrelations. In bottleneck-order solutions, projects are ranked based on their link volume-capacity ratios, which measure their congestion severity. This assumes that more congested links should have higher priority for improvement.

The fitness function is equal to the value of the objective function which, as stated earlier, is computed through the traffic assignment model. In maximization problems, the selection probability corresponds to the value of the objective function. In minimization problems the selection probability correlates inversely with the objective function value. To avoid prematurity properties, a ranking method proposed by Michalewicz (1995) is used. In this method the population is ordered from best to worst. Then, based the exponential ranking value, the selection probability of each chromosome is assigned, assuming the lowest fitness value is one (Michalewicz, 1995). Letting q be the selective pressure$\in$ [0,1], the selection probability is defined as follows:

$$P_i = c * q(1 - q)^{i-1}, \quad c = 1/[1 - (1 - q)^{PopSize}] \tag{4}$$

Next, a roulette wheel approach is used to choose appropriate parents based on their selection probabilities (Michalewicz, 1995). This process is conducted by spinning the roulette wheel once for each individual in the population. Each time a random number r [0,1] is generated, the i_{th} chromosome is selected so that $w_{i-1} < r \leq w_i$, where w_i is the cumulative probability for each chromosome. Then the crossover and mutation operators are applied to reproduce offspring and create the new population. Common methods of mutation and crossover are fairly inefficient for sequencing problems since they construct many infeasible solutions with repetitive project numbers within one sequence. To avoid producing such solutions, some other genetic operators are employed to solve the project sequencing problem. These operators, adapted from Wang (2001), include Partial Mapped Crossover (PMX), Position Based Crossover (PBX), Order Crossover (OX), Order Based Crossover (OBX), Edge Recombination Crossover (ERX), Insertion Mutation (IM), Inversion Mutation (VM) and Reciprocal Exchange Mutation (EM).

6. ANALYSIS FRAMEWORK

The framework of the general proposed method for selecting, sequencing and scheduling interrelated road projects is presented in Figure 1. The proposed combination of traffic assignment and metaheuristic algorithms may be used to evaluate any sequence of projects and find a near-optimal solution to the project selection and scheduling problem.

The pseudo algorithm provided in this section explains step-by-step how this problem is tackled. First, the traffic assignment algorithm known as Frank-Wolfe, which is also used in this study to evaluate the system at various stages, is described. This user equilibrium model distributes flow in the network in a way that no individual user can reduce its trip cost by switching routes. The second part describes the optimization algorithm. It also explains how the user equilibrium algorithm is used within the GA to evaluate the objective function i.e. fitness value of the population. In this case, each chromosome presents a string of numbers which is the sequence of projects. The fitness value i.e. the objective function for each chromosome is estimated by re-running the user equilibrium model at relatively short intervals during the analysis period, and thereby estimate the effects of additional projects on traffic volumes and speeds throughout the system.. This in fact captures the interrelation

among projects. Equation 1 yields the present value of total cost which is also the fitness value for the chromosome. Accordingly, new generations are created and evaluated until the GA's termination condition is met.

Evaluation Model – User Equilibrium (Frank-Wolfe)

Given a current travel time for link a, t_a^{n-1} the nth iteration of the convex combination algorithm is summarized as follows:

1. Initialization: all or nothing assignment assuming t_a^{n-1} which yields x_a^n.
2. Updating travel time: use a BPR function $t_a^n = t_a(x_a^n) = t_0(1 + 0.15 \left(\frac{v}{c}\right)^4)$.
3. Direction finding:
    - Find shortest paths using Dijkstra Algorithm based on t_a^n
    - All or nothing assignment considering t_a^n which yields auxiliary flow y_a^n.
4. Line search: find α that solves $min \sum_a \int_0^{x_a^n + \alpha(y_a^n - x_a^n)} t_a(\omega)d\omega$.
5. Move: set $x_a^{n+1} = x_a^n + \alpha_n(y_a^n - x_a^n)$, $\forall \alpha$.
6. Convergence test: If a convergence criterion met, stop. Otherwise set n=n+1 and go to step 1.

Optimization Model – Genetic Algorithm

1. $t \leftarrow 0$
2. Initial population: Set initial population [P(t)].
3. Evaluate population:
    - For each chromosome (sequence of projects), run User Equilibrium after each project (gene) is implemented.
    - Obtain travel time w_{ij}, volume, VMT.
    - Compute the fitness value through eq.1.
4. **While** not termination, **do**
    - Select parents [$P_p(t)$]
    - Reproduce offspring by crossover operators [$P_c(t)$]← [$P_p(t)$]
    - Mutate [$P_c(t)$]
    - Create next generation [P(t+1)]
    - $t \leftarrow t+1$

End.

5. Obtain optimized sequence of projects.

7. CASE STUDY

In the literature, simple examples of related problems have been published, e.g. by Tao and Schonfeld (2006). A more complex example, namely the Sioux Falls network (LeBlank et al., 1975) is used as a case study here. Sioux Falls is the largest city in the U.S. state of South Dakota. Its simplified network with 24 nodes and 76 links, shown in Figure 3, is used here for testing purposes. It is assumed for this example that the demand grows exponentially over the planning horizon:

$$d_{ij}^t = d_{ij}^0 * (1 + r)^t \tag{5}$$

where d_{ij}^t is the demand between origin i and destination j, d_{ij}^0 is the base demand for the ij origin and destination (O/D) pair at time 0, and r is the growth rate per period.

After running the traffic assignment model, the critical lanes with high volume-capacity ratios are selected as an initial set of candidate projects. Our model allows volume-capacity ratios above 1.0 since we use a BPR function for estimating link performances. Since the demand matrix is symmetric for O/D pairs, each link expansion improvement is assumed to be implemented in both directions between the two connected nodes, i.e. each project is defined as expanding two links between a pair of connected nodes. This assumption is also justified economically because it saves costs in using mobilized construction equipment and other resources. To find appropriate initial solutions, the traffic assignment model is run for all improvement scenarios. The first column in Table 1 shows the sequence of projects ranked by their benefit-cost ratio in descending order. In this context, the benefit is the present value of travel time savings, and the cost is the present value of implementation cost (greedy order solution). The third column displays the sequence of projects based on their congestion severity, where links with lower service levels have higher priorities (bottleneck order solution).

TABLE 1 Greedy Order and Bottleneck Order Solutions

Greedy Order Solution (Link #)	Project Benefit (dollar)	Bottleneck Order Solution (Link #)	V/C Ratio
11	$217,300,346	11	2.17
36	$193,368,891	36	1.89
3	$189,404,178	34	1.79
12	$161,423,613	14	1.62
9	$117,425,401	9	1.59
15	$91,362,677	27	1.48
2	$87,751,583	35	1.42
25	$71,863,522	12	1.41
21	$70,811,860	15	1.36
4	$69,331,975	21	1.35
27	$68,775,533	3	1.35
37	$61,764,580	13	1.32
16	$61,099,054	30	1.31
22	$60,702,083	37	1.22
13	$60,135,953	22	1.21
14	$59,110,008	4	1.11
35	$44,182,898	2	1.11
30	$36,073,907	16	1.09
34	$5,242,573	25	1.04

After identifying an initial set of candidates, all projects are further investigated through a benefit-cost analysis to identify and rank the initial economically beneficial projects. It is assumed that each improvement project adds one lane, which is equivalent to 700 vehicles/hour additional capacity to each link, and the equivalent annual cost of each lane

expansion is assumed to be 4,000,000 $/lane-mile (Zhang et al., 2013). The main cost saving of link expansion projects is the reduced travel time for all the users. These travel time reductions can be computed through the traffic assignment model by comparing the total system travel time before and after project implementation. Next, the previously described GA is used to find near-optimal solutions for the sequence and schedule of selected projects. When optimizing, we seek a sequence of projects which can be implemented within the planning horizon (30 years). Therefore, every project with a scheduled completion time beyond the planning horizon is eliminated from the sequence.

8. RESULTS

As discussed previously, a traffic assignment model is used to evaluate the candidate projects over the planning horizon and a GA is used to find near-optimal solutions. This section analyzes the GA results and compares the basic scenario without improvement projects to the scenarios with implemented projects.

TABLE 2 Optimal Sequence and Schedule

Optimal Sequence	Completion Time (year)
11	1.8
34	5.9
36	8.8
9	10.8
14	14.8
3	16.2
35	20.7
27	22.7
37	25.0
12	28.0
NPV of Total Cost$\times 10^6$($)	8535.93

In this analysis the average GA running time per iteration is 300 sec and the entire analysis takes about 8 hours to run.

Table 2 presents the optimal sequence and the corresponding schedule of projects along with the objective value. The first column presents the link identifiers as ordered in the optimized solution. As stated earlier, each link expansion improvement is assumed to be implemented in both directions between the two connected nodes. Accordingly, the optimized schedule is directly determined by the sequence of selected projects, assuming it is efficient to fund and finish one project at a time, and gain its benefits as soon as it is completed. Thus, as

explained in section 2, successive projects in the sequence are completed when the available cumulative budget equals the cumulative project cost. Figure 5 shows the accumulated total delay costs for three scenarios: (i) no project implementation, (ii) project implementation based on greedy solution, and (iii) optimized project schedule. These results indicate that at the end of 30 years, the improvement projects can save up to 21% of the total delay costs compared to no project implementation and 10.5% compared to the greedy order solution.

In addition to Sioux Falls network which is fairly small, this method is also applied to the much larger Anaheim network, which is displayed in Figure 5. It has 416 nodes (of which 38 are origin/destination centroids), 914 links, and 1406 O-D pairs. All the network-related information is extracted from (Bar-Gera, 2011). In this case, we tested the algorithm for 20, 40, 80 and 100 candidate projects. Table 3 compares CPU times for the Anaheim and Sioux Falls networks. It can be seen that a larger network significantly increases the CPU time. The results also indicate that the network size affects the CPU time much more than the number of projects. In this case, where number of links in the Anaheim network is 12 times higher, the CPU time per generation becomes almost 115 times higher. This occurs because the traffic assignment algorithm has to evaluate the entire network regardless of the number of projects. Also, the number of generations for comparable precision is likely to increase with network size. In conclusion, this method is applicable to fairly large networks with numerous projects, but computational improvements would be desirable for analyzing very large networks.

Table 3 CPU Time per Generation (Sec)

Sioux Falls	Number of projects	5	10	15	20
	CPU time	51.65	91.26	149.25	161.53
Anaheim	Number of projects	20	40	80	100
	CPU time	10,472	12,764	16,897	18,533

9. ALGORITHM TESTING

To evaluate the results emerging from this algorithm, an exhaustive enumeration is carried out for the Sioux Falls network. Since the enumeration of the original problem with 20 candidate projects (i.e. 20! possible solutions) is lengthy and requires extensive computation time, this test is done for smaller problems with fewer projects. In this case, we consider four problems with 4, 5, 6 or 7 projects to be ranked. Each case is solved both by the GA and by a complete enumeration which evaluates each possible combination of projects and renders the exact solution. The results presented in TABLE 4 indicate that the GA yields the exact solution from enumeration in all four cases.

Table 4 Complete Enumeration Test

Number of projects	Solution space	Complete enumeration		GA solution	
		Total system cost $* 10^6$	Optimal sequence	Total system cost $* 10^6$	Optimal sequence
4	4!=24	90980	3,2,1,4	90980	3,2,1,4
5	5!=120	94248	3,2,5,4,1	94248	3,2,5,4,1

| 6 | 6!=720 | 98009 | 3,2,5,4,1,6 | 98009 | 3,2,5,4,1,6 |
| 7 | 7!=5040 | 99301 | 3,2,5,4,1,6,7 | 99301 | 3,2,5,4,1,6,7 |

In general, it is impractical to fully guarantee that the results of heuristic algorithms are globally optimal, and it is somewhat difficult to assess the goodness of solutions obtained by the evolutionary methods. In this study, a statistical experiment is conducted to examine the effectiveness of the algorithm. For this purpose, first a sample of randomly generated independent solutions is created. The next step is to fit an appropriate distribution to the fitness values. The final step is to calculate the cumulative probability of the solution found by the algorithm based on the fitted distribution. It is desirable to obtain a very low probability to demonstrate the goodness of the solution. Accordingly, a random sample of 50,000 solutions is created, for which the objective function minimum is 8709.19×10^6 and maximum is 15769.69×10^6. After exploring different distributions, the Lognormal (mu= 9660, sigma= 0.0248) distribution is found to yield the best fit. Figure 6 shows the fitted distribution and the data derived from random sampling. It is evident that the minimum value in the distribution of 50,000 random solutions is higher (costlier) than the optimal solution presented in TABLE 2. In other words, the solution found by the algorithm excels all the random solutions in the distribution.

The cumulative probability of the best solution found by the GA according to the Lognormal distribution is $p = F(x|\mu, \sigma) = F(8535.93 \times 10^6 | 9660, 0.0248) = 3.597 \times 10^{-5}$ which can be derived from the following equation:

$$p = F(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{e^{\frac{-(\ln(t)-\mu)^2}{2\sigma^2}}}{t} dt \qquad (6)$$

This result implies that the best solution obtained by the algorithm dominates 99.999% of the random solutions in the distribution. Therefore, the solution found by the GA, although not guaranteed to be globally optimal, is very good compared to other possible alternatives in the solution space and the deviation from global optimality is likely to be very small compared to uncertainties and errors in the problem's inputs.

10. CONCLUSIONS

The capacity expansion of links in road networks is a typical example of interrelated alternatives for which the selection and sequencing of projects becomes a challenging optimization problem with a "noisy" objective function surface. Common methods for evaluating and prioritizing such problems are often incapable of capturing the interactions among projects, and are mostly limited to pair-wise or at best n-wise interactions. The main contribution of this study is to demonstrate how a traffic assignment model can be combined effectively with a GA in a multi-period analysis for planning and prioritizing purposes while capturing interactions among projects. We also design the algorithm to account for the possibility that candidate projects may become economically justified or unjustified after the implementation of previous projects. Another contribution is to reformulate the budget constraint to include possible internal funding from fuel taxes. Also, we assume that the demand changes during the planning horizon (growing exponentially in our example).

Finally, we demonstrate this methodology by conducting a case study and present a statistical test of the goodness of the heuristic results.

In this study, a GA approach is employed here to optimize the selection and scheduling of link expansion projects. The study uses a simple traffic assignment model to evaluate the objective function and combines it with the GA to optimize the solution. Although road expansion projects are the focus of this study, the proposed methodology should be applicable to general cases involving more complex systems. More specifically, GAs can optimize very intractable objective functions without requiring restrictive assumptions about their structures. This allows analysts to effectively combine an appropriate evaluation tool (e.g. microscopic simulation, simulation approximates, queuing or neural networks, depending on the problem) with the GA, and to solve the planning and scheduling problem for a variety of interrelated alternatives.

Future research may focus on developing general frameworks for solving the problem of planning and prioritizing interrelated alternatives in a wide range of applications. Although many components of such a general method exist, they could benefit from further improvements. Accordingly, the work presented in this paper may be extended by incorporating more complex evaluation models (e.g. micro simulation) to capture saturation effects in networks. Future work may also account for uncertainties of important variables, and consider other possibilities, such as multiple alternatives per location, facility changes over time at the same location, and traffic delays during construction. Computational improvements in the algorithm would be desirable, e.g. by distributing GA's operators among multiple computer processors. It may also be interesting to optimize particular projects endogenously instead of selecting them from among pre-specified projects.


11. ACKNOWLEDGEMENTS

12. REFERENCES

1. Bagloee, S. A., and Tavana, M. (2012) An Efficient Hybrid Heuristic Method for Prioritising Large Transportation Projects with Interdependent Activities. *International Journal of Logistics Systems and Management*, Vol.11, No. 1, pp.114-142.
2. Bagloee, S. A., and Asadi, M. (2015). Prioritizing road extension projects with interdependent benefits under time constraint. *Transportation Research Part A: Policy and Practice*, 75, pp.196-216.
3. Balamurugan, K., Selladurai, V., and Ilamathi, B. (2006). Solving unequal area facility layout problems using genetic algorithm. International Journal of Logistics Systems and Management, 2(3), 281-301.
4. Beenhakker, H. L., and Narayanan, V. (1975). Algorithms for scheduling projects with limited resources. *The Engineering Economist*, 21(2), pp.119-140.
5. Bhattacharyya, R., Kumar, P., and Kar, S. (2011) Fuzzy R&D Portfolio Selection of Interdependent Projects. *Computers & Mathematics with Applications*, Vol. 62, No. 10, pp. 3857-3870.
6. Bouleimen, K., and Lecocq, H. (2003) A New Efficient Simulated Annealing Algorithm for the Resource-Constrained Project Scheduling Problem and Its Multiple Mode Version. *European Journal of Operational Research*, Vol. 149, No. 2, pp. 268-281.
7. Chen, X., Zhu, Z., He, X., and Zhang, L. (2015) Surrogate-based Optimization for Solving Mixed Integer Network Design Problem. *Transportation Research Record: Journal of the Transportation Research Board, No. 15-4556.*
8. Chiu, C. Y., and Park, C. S. (1998). Capital budgeting decisions with fuzzy projects. *The Engineering Economist*, 43(2), pp. 125-150.
9. Dai, D.M., and Schonfeld, P. (1998) Metamodels for Estimating Delays through Series of Waterway Queues. *Transportation Research Part B: Methodological*, Vol. 32, No.1, 1998, pp. 1-19.
10. Dickinson, M. W., Thornton, A. C., and Graves, S. (2001). Technology portfolio management: optimizing interdependent projects over multiple time periods. *Engineering Management, IEEE Transactions on*, 48(4), pp.518-527.
11. Disatnik, D. J., and Benninga, S. (2007) Shrinking the Covariance Matrix. *The Journal of Portfolio Management*, Vol. 33, No. 4, pp. 55-63.
12. Dueñas-Osorio, L., Craig, J. I., Goodno, B. J., and Bostrom, A. (2007) Interdependent Response of Networked Systems. *Journal of Infrastructure Systems*, Vol.13, No. 3, pp. 185-194.
13. Durango-Cohen, P. L., and Sarutipand, P. (2007) Capturing Interdependencies and Heterogeneity in the Management of Multifacility Transportation Infrastructure Systems. *Journal of Infrastructure Systems*, Vol.13, No. 2, pp. 115-123.
14. Cruz, L., Fernandez, E., Gomez, C., Rivera, G., & Perez, F. (2014). Many-objective portfolio optimization of interdependent projects with 'a priori'incorporation of decision-maker preferences. Appl. Math, 8(4), 1517-1531.
15. Frank, M., and Wolfe, P. (1956) An Algorithm for Quadratic Programming. *Naval Research Logistics Quarterly*, Vol. 3, No. 1, pp. 95-110.
16. Golberg, D. E. (1989) Genetic algorithms in Search, Optimizaion, and Machine Learning. *Addion Wesley*.
17. Haq, A. N., and Kannan, G. (2006). Two-echelon distribution-inventory supply chain model for the bread industry using genetic algorithm. International Journal of Logistics Systems and Management, 2(2), 177-193.
18. Jong, J. C., and Schonfeld, P. (2001) Genetic Algorithm for Selecting and Scheduling Interdependent Projects. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, Vol.127, No. 1, pp. 45-52.
19. Koc, A., Morton, D. P., Popova, E., Hess, S. M., Kee, E., and Richards, D. (2009). Prioritizing project selection. *The Engineering Economist*, 54(4), pp.267-297.

20. LeBlanc, L. J., Morlok, E. K., and Pierskalla, W. P. (1975) An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem. *Transportation Research*, Vol.9, No. 5, pp. 309-318.
21. Li, Z., Roshandeh, A. M., Zhou, B., and Lee, S. H. (2013) Optimal Decision Making of Interdependent Tollway Capital Investments Incorporating Risk and Uncertainty. *Journal of Transportation Engineering*, Vol. 139, No.7, pp. 686-696.
22. Li, X., Fang, S. C., Guo, X., Deng, Z., and Qi, J. (2016). An extended model for project portfolio selection with project divisibility and interdependency. Journal of Systems Science and Systems Engineering, 25(1), 119-138.
23. Markowitz, H. (1952) Portfolio Selection. *The Journal of Finance*, Vol.7, No.1, , pp. 77-91.
24. Michalewicz, Z., and Janikow, C. Z. (1991) Genetic Algorithms for Numerical Optimization. *Statistics and Computing*, Vol. 1, No. 2, pp. 75-91.
25. Michalewicz, Z. (1995) Genetic algorithms + data Structure = evolution programs. *Springer*.
26. Rebiasz, B., Gaweł, B., and Skalna, I. (2014). Capital budgeting of interdependent projects with fuzziness and randomness. Information systems architecture and technology, Wrocław, Oficyna Wydawnicza Politechniki Wrocławskiej, 125-135.
27. Sandhu, M. (2006). Project logistics with the dependency structure matrix approach–an analysis of a power plant delivery. International Journal of Logistics Systems and Management, 2(4), 387-403.
28. Shayanfar, E., Abianeh, A. S., Schonfeld, P., and Zhang, L. (2016) Prioritizing Interrelated Road Projects Using Meta-Heuristics. *Journal of Infrastructure Systems*, 04016004.
29. Szimba, E., and Rothengatter, W. (2012) Spending Scarce Funds More Efficiently—Including the Pattern of Interdependence in Cost-Benefit Analysis. *Journal of Infrastructure Systems*, Vol.18, No. 4, pp. 242-251.
30. Tao, X., and Schonfeld, P. (2005) Lagrangian Relaxation Heuristic for Selecting Interdependent Transportation Projects Under Cost Uncertainty. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1931*, pp. 74-80.
31. Tao, X., & Schonfeld, P. (2006). Selection and scheduling of interdependent transportation projects with island models. *Transportation Research Record: Journal of the Transportation Research Board*, (1981), 133-141.
32. Tao, X., and Schonfeld, P. (2007) Island Models for a Stochastic Problem of Transportation Project Selection and Scheduling. *Transportation Research Record: Journal of the Transportation Research Board, No.* 2039, pp.16-23.
33. Tofighian, A. A., and Naderi, B. (2015). Modeling and solving the project selection and scheduling. Computers & Industrial Engineering, 83, 30-38.
34. Wang, S. L., and Schonfeld, P. (2008) Scheduling of Waterway Projects with Complex Interrelations. *Transportation Research Record: Journal of the Transportation Research Board, No. 2062*, pp. 59-65.
35. Wang, S. L., and Schonfeld, P. (2005) Scheduling Interdependent Waterway Projects through Simulation and Genetic Optimization. *Journal of Waterway, Port, Coastal, and Ocean Engineering,* Vol.131, No. 3, pp. 89-97.
36. Wei, C.H., and Schonfeld, P. (1994) Multi-period Network Improvement Model. *Transportation Research Record: Journal of the Transportation Research Board, No. 1443*, pp. 110-118.
37. Wang, S. L. (2001) Simulation and Optimization of Interdependent Waterway Improvement Projects. PhD dissertation, Univ. of Maryland, College Park, MD.
38. Zhang, L., Ji, M., and Ferrari, N. (2013) Comprehensive Highway Corridor Planning with Sustainability Indicators (Final Report), Maryland State Highway Administration.
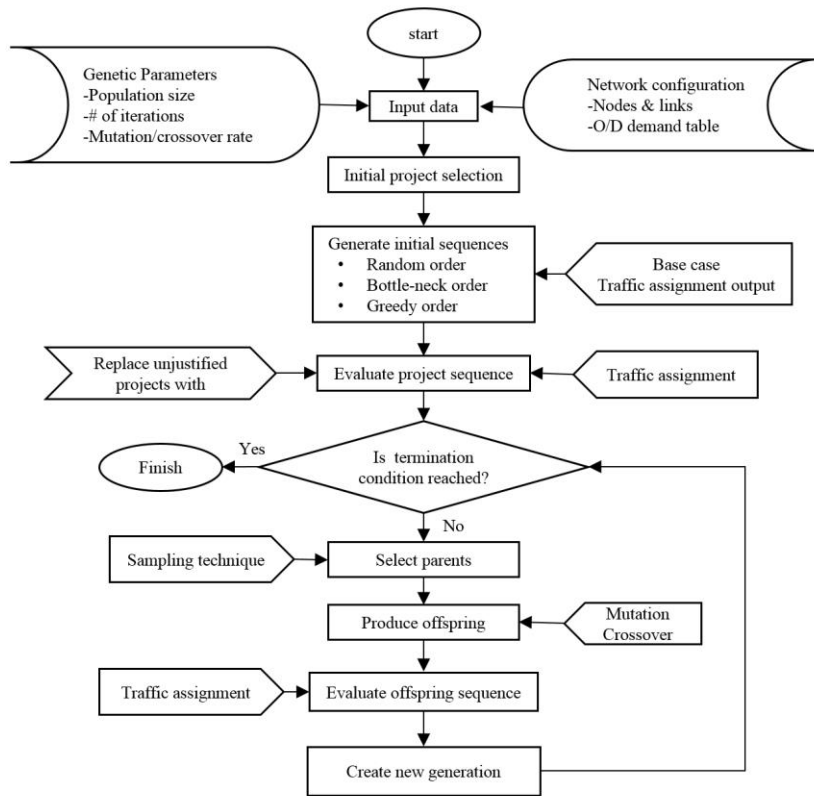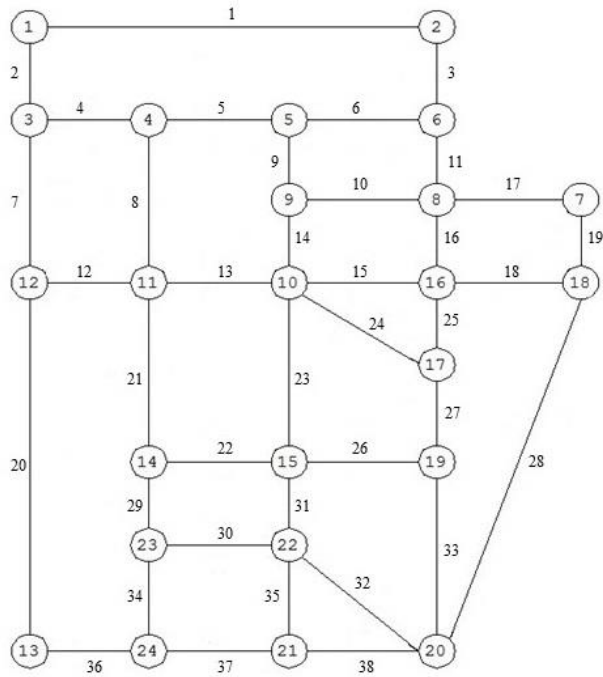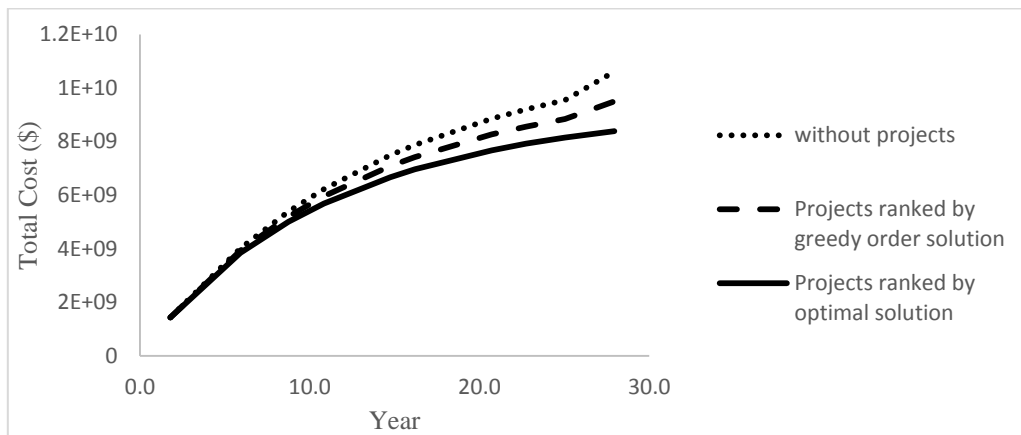
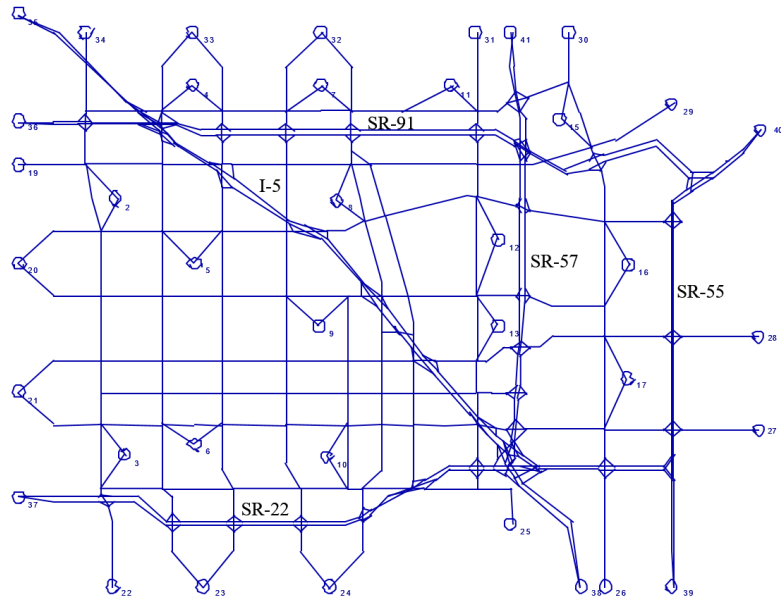**FIGURE 1 Framework of Optimization Process.**
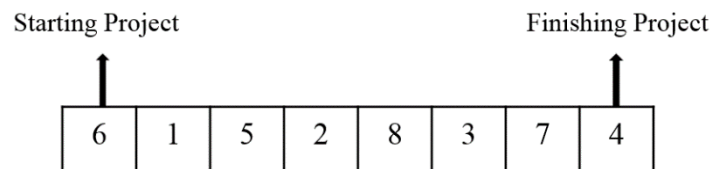
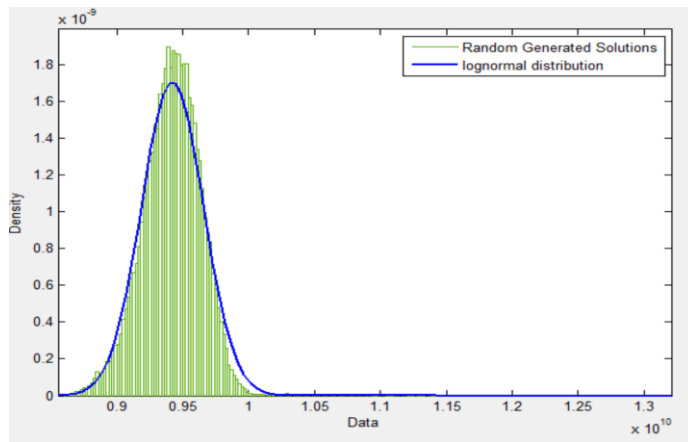**FIGURE 2 Example of a Feasible Solution.**

FIGURE 3 Sioux Falls Network.

FIGURE 4 Accumulated Total Delay Cost with and without projects.

Figure 5 Anaheim Network.

FIGURE 6 Fitted Lognormal Distribution.

Appendix 2 - Jovanovic, U., Shayanfar, E. and Schonfeld, P. "Selecting and Scheduling Link and Intersection Improvements in Urban Networks," *Transportation Research Record,* v2672, Dec. 2018, pages 1-11.

Selecting and Scheduling Link and Intersection Improvements in Urban Networks

Uros Jovanovic, Elham Shayanfar, and Paul M. Schonfeld

Department of Civil and Environmental Engineering, University of Maryland, College Park, MD

Corresponding Author:

Address correspondence to Paul M. Schonfeld: pschon@umd.edu

Abstract

Deciding which projects, alternatives or investments to implement is a complex and important problem not only in transportation engineering, but in management, operations research and economics. Projects are interrelated if their benefits or costs depend on which other projects are implemented. Furthermore, in the network development problem analyzed here, the timing of projects also affects the benefits and costs of other projects. This paper presents a method for optimizing the selection and scheduling of interrelated improvements in road networks that explicitly considers intersections. The Frank Wolfe algorithm, which is modified here to consider intersections, is used for evaluating network improvements as well as for traffic assignment. Intersections are modelled with pseudo-links whose delays are estimated with Akcelik's generalized model. The objective is to minimize the present value of total costs (including user time) by determining which projects should be selected and when they should be completed. A genetic algorithm is used for optimizing the sequence and schedule of projects.

For decades transportation engineers have been dealing with the problem of evaluating, selecting and scheduling infrastructure projects. Considered alternatives can be classified as follows:

- Mutually exclusive: Only one alternative may be selected;

- Independent: The benefits and costs of alternatives

are independent of which alternatives are selected or when those are implemented;

- Interdependent (interrelated).

Interrelated alternatives pervade transportation net- works since improvements alter the flows, and hence bene- fits, on other network components. This paper aims to show how a traffic assignment model can be used to evaluate the objective function of an investment planning optimization problem for an urban road network, especially by showing how intersections can be included in the traffic assignment. A method is presented for evaluating, selecting and scheduling interdependent improvement alternatives in urban road networks, which extends Shayanfar et al. (1) by considering intersection improvements in addition to link widening alternatives. It is shown how a traffic assignment model can be effectively modified to consider intersection flows and delays by introducing pseudo-links. Adding pseudo-links for each of three movements (left, through and right) at each approach of a four-leg intersection, creates a total of 12 pseudo-links per intersection. Moreover, a traffic assignment model is shown to be effectively combined with a genetic algorithm for planning and prioritizing purposes while considering interrelations among candidate projects. The background section reviews some prior studies on intersection delay, selection and scheduling of project alternatives, and traffic assignment. The next two sections present the evaluation model and the genetic algorithm used for optimizing the project selection and schedule. A case study is presented on the Sioux Falls network and the results obtained with the modified traffic assignment model and genetic algorithm in optimizing the network development schedule. Conclusions and suggestions for extensions are presented in the last section.

Background

Intersections are crucial components in urban road net- works since they affect traffic capacity and delay at least as much as road links. Typically, four-leg intersections allow up to 12 legal vehicular movements and 4 legal pedestrian crossing movements. Traffic signals assign right-of-way, and can significantly reduce the number of conflicts, thus regulating the traffic flow. One of the many disadvantages of traffic signals is the possibility of excessive delay which can congest the network, which, in turn, increases cost, pollution and driver anxiety. Early studies on delays at signalized intersections include Wardrop (2), who assumed that vehicles enter intersections with uniform arrivals, and Webster (3) who studied delays for vehicles at pretimed signals and optimized their settings.

Delay relates to the amount of excess travel time, fuel consumption, and the frustration and discomfort of drivers. Delay can also be used to compare the performance of an intersection under different demand, control and operating conditions. For intersections, delay can be calculated simply, as the difference in the departure time and the arrival time of a vehicle. Estimation of overflow delay is one of the major difficulties in developing delay models at signalized intersections. The difficulty is obtaining a simple and easily computable formula for overflow delay and has forced researchers and analysts to search for approximations and boundary values. Numerous intersection delay models have been developed, including Webster's (3), Highway Capacity Manual (HCM) (4), Australian (variation

of the Akcelik delay model (5, 6), and Canadian (7). The delay model used here is Akcelik's, because it gives delay values close to the HCM formula for v/c \ 1.0, but with fewer assumptions about parameters. It is expressed in Equations 1 and 2 as

$$d = 0.5\frac{C(1-\lambda)^2}{(1-\lambda x)}$$
$$+ 900Tx^n\left[(x-1) + \sqrt{(x-1)^2 + \frac{m(x-x_0)}{cT}}\right] \quad (1)$$

and

$$x_0 = a + bsg \quad (2)$$

where

d = average overall delay (sec/veh),

C = cycle time (sec),

l = fraction of the cycle which is effectively green for the phase under consideration,

x = v/c ratio,

T = flow period (h),

c = link capacity (veh/h),

m, n, a, b = calibration parameters, whose values are available for different delay models (e.g., Australian, Canadian, TRANSYT (8), and HCM) in Akcelik's paper (5), and

s*g = capacity per cycle (veh/cycle).

Parameters n, m, a and b according to Akcelik's papers (5, 6) have the following values respectively: 0, 8, 0.5, and 0. Therefore, the two equations above become

$$d = 0.5\frac{C(1-\lambda)^2}{(1-\lambda x)}$$
$$+ 900T\left[(x-1) + \sqrt{(x-1)^2 + \frac{8(x-0.5)}{cT}}\right] \quad (3)$$
$$x_0 = 0.5$$

The overall delay d_I at an intersection can be calculated as

$$d_I = \frac{\sum d_A v_A}{\sum v_A} \quad (4)$$

where d_A is delay on approach A, and v_A is volume on approach A. Heidemann (9) and Olszewski (10) used probability distribution function to estimate delay at signalized intersections.In their models, the probability distributions of delay were obtained from the probabilities of queue lengths.

Among many approaches used to tackle the problem of project selection and scheduling are integer programming, used by Weingartner (11) and by Cochran et al. (12), and dynamic programming, used by Weingartner (11) and by Nemhauser and Ulman (13). One notable study on interrelated projects is Weingartner's (11) which presents, among other problems, interdependent projects with budget constraints.

Mehrez et al. (14) use a multi-attribute function to specify the decision maker's preference with a zero-one budget model to solve the problem of selection of interrelated multi-objective long-range projects. The authors define a set of n indivisible projects contributing to m tangible and intangible attributes with L limited resources available for T periods. In addition, they use a utility function with m attributes and regard each project as a collection of subprojects, each one contributing to one of the attributes affected by the projects.

Evaluation Model

Traffic assignment can be formulated as the problem of finding the equilibrium flow pattern over a given transportation network, if its graph representation, the associated link performance function and an origin–destination (O-D) matrix is known. Assignment of traffic flows on network links is a result of equalizing transportation demand (O-D matrix) and transportation supply (link and node capacity, management actions). A reasonable assumption is that all travelers try to minimize their own travel time between their own origins and destinations. Other assumptions are that travel times increase with link flows, and all individuals behave identically. User equilibrium (stable condition) is achieved when no traveler can improve their travel time by changing route. Notable publications that dealt with traffic assignment include Florian (15), Sheffi (16), and Boyce and Ran (17, 18). However, none of these consider intersection characteristics and performance.

This paper applies the convex combination algorithm developed by Frank and Wolfe (19) to evaluate link and intersection expansion projects upon their implementation in the network. The Frank-Wolfe (FW) algorithm is an iterative algorithm used for solving a user equilibrium traffic assignment which is a nonlinear programming problem with convex objective function and linear constraints. Given t_a^0 a (initial travel time for link a), the convex combination algorithm is as follows:

1. Set counter $n = 1$,
2. Initialization: Perform all or nothing assignment assuming t_a^{n-1}, which yields flows x_a^n
3. Update: Set link travel time (Bureau of Public Road (BPR) function) $t_a^n = t_a(x_a^n) = t_0\left(1 + 0.15\left(\frac{v}{c}\right)^4\right)$
4. Direction finding: Perform all-or-nothing assignment based on $\{t_a^n\}$, which will yield a set of (auxiliary) flows $\{y_a^n\}$
5. Line search: find $\propto_n$ that solves the following problem:

$$\max_{0 \leq x \leq 1} \sum_a \int_0^{x_a^n + \alpha\left(y_a^n - x_a^n\right)} t_a(\omega)d\omega \qquad (5)$$

6. Move: Set $x_a^{n+1} = \alpha\left(y_a^n - x_a^n\right), \forall \alpha$
7. Convergence test: If a convergence criterion is met, stop. Otherwise, set $n = n + 1$ and return to step 2.

Problem Statement

The problem considered here is NP hard (20) with a nonconvex objective function. The problem grows rapidly as the number of candidate projects increases, and can be classified as a combinatorial optimization problem. This type of problem involves finding values for discrete variables in such a way that the optimal solution is found with respect to the objective function. Many practical problems can be classified as combinatorial optimization problems such as the shortest path algorithm. Other examples are the optimal assignment of employees to tasks to be performed and the traveling salesman problem. Dorigo et al. (21) formulated a combinatorial optimization problem U as a triple (S, f, O), where S is the set of candidate solutions (sequence of projects), f is the objective function (present value of total costs) which assigns an objective function value f (s) to each candidate solution s 2 S, and O is the set of constraints (budget constraint in our case). The solutions belonging to the set ~S S of candidate solutions that satisfy the constraints O are called feasible solutions. The goal, according to Dorigo et al. (21), is to find a globally optimal feasible solution s* (optimal sequence of projects).

In this study, the present value of total cost during the analysis period is the objective function, subject to a budget constraint. The total cost consists of: (i) supplier cost, defined as the present value of all project costs, and (ii) user cost, defined as the delay multiplied by the value of time. Accordingly, the objective function can be formulated as

$$Z = \sum_{j=1}^{T} \frac{v}{(1+r)^j} \sum_{i=1}^{n_l} w_{ij} + \sum_{j=1}^{T} \frac{1}{(1+r)^j} \sum_{i=1}^{n_{pl}} c_i x_i(t)$$
$$+ \sum_{j=1}^{T} \frac{v}{(1+r)^j} \sum_{i=1}^{n_l} d_{ij} + \sum_{j=1}^{T} \frac{1}{(1+r)^j} \sum_{i=1}^{n_{pl}} C_i X_i(t) \quad (6)$$

where

w_{ij} = waiting time on link i in year j,

c_i = present value of the cost of link project i,

n_{pl} = number of link projects (link improvements),

n_l = total number of links,

n_I = total number of intersections,

n_{pI} = number of intersection improvement projects,

C_i = present value of the cost of intersection project i,

v = value of time, and

r = interest rate.

The cost of intersection project i can be written as

$$C_i = C_{c_i} + C_{p_i} = A_{I_i} \cdot 51.6 + A_i \cdot 20 \quad (7)$$

where

C_{ci} = capital cost of improvement of intersection i ($/ft2),

C_{pi} = cost of pavement maintenance of intersection I ($/ft2),

A_{Ii} = area of the land needed to improve intersection I (ft2),

A_i = overall area of the intersection i (ft2)

The objective function is bound by the following cumulative budget constraint (22) as

$$\sum_{i=1}^{n_p} c_i x_i(t) \le \int_0^t B(t)dt, 0 \le t \le T \qquad (8)$$

$$\begin{cases} x_i(t) = 0 \ \ if \ \ t < t_i \\ x_i(t) = 1 \ \ if \ \ t > t_i \end{cases}$$

where ti is the time when project i is finished, and $x_i(t)$ is a binary variable specifying whether project i is finished by time t. Since in most realistic problems the cumulative budget constraint is binding, that is there is never enough funding for all the available projects that are worth implementing, the optimized project sequence represented by the set of all tis uniquely determines the schedule of projects (1, 22).

Optimization Method

A genetic algorithm (GA) is a search technique inspired by biologic natural selection and evolution: ''survival of the fittest''. Traditional techniques evaluate only one potential solution at a time when searching for the optimal solution, while a GA searches by concurrently examining a population of solutions. First, the GA generates many different solutions and computes their fitness value (which in most cases is the objective function value). Then, solutions are ranked based on their fitness value. Solutions with better fitness values are saved, while others are discarded. Some saved solutions are chosen as parents, and genetic operators, such as mutation and recombination operators, are applied on them to create a new generation of solutions. This process is repeated, until the specified number of generations is achieved or until the fitness function stops improving significantly. The GA includes the following steps (23):

1. Code the problem and determine the values of the parameters.

2. Form an initial population which contains n strings, where n depends on the type of problem examined. Evaluate the fitness function of every string.

3. Assuming the probability of choice is proportional to values of fitness function, choose n potential parents.

4. Randomly choose two or more parents and apply operators such as recombination and mutation

operators to create offspring until a new population of n offspring is created.

5. Evaluate the fitness function for the new population for every offspring.

6. If the stopping criterion is reached, terminate the algorithm, and report the optimized solution (one

with the best fitness value). Otherwise, return to step 3.

In this study, the initial population of the GA is generated randomly and solutions are represented by integer digits showing the sequence of the projects being implemented. Each individual in a population is defined as a string of numbers, each corresponding to a specific project in a sequence. The fitness function is the value of the objective function and is computed through the traffic assignment model.

Table 1. Input Parameters

Parameter	Definition	Value and unit
c_i	Cost of lane addition	$3 million/lane.mile
g	Demand growth rate	0.01/year
Ln_w	Lane widths	11 ft
B	Available budget	$1.5 million/year
v	Value of time	$15/veh.hr
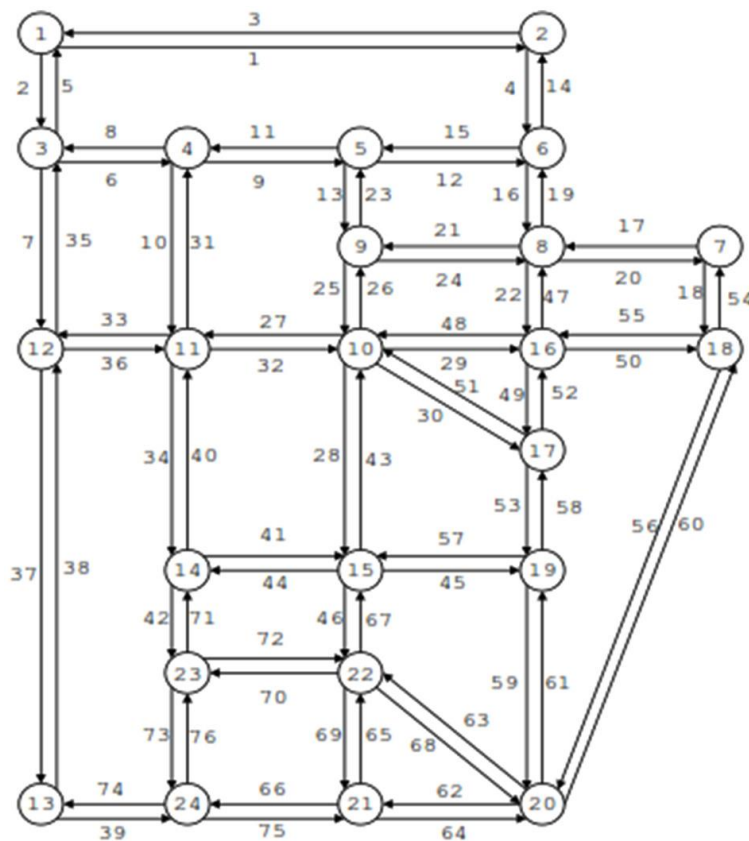
Figure 1. Graphical representation of the Sioux Falls network.

Case Study

The Sioux Falls network adopted from LeBlanc et al. (24) is used here as a case study. This network differs from the real network since it mainly includes the city's major arterials. It has been used in many previous studies. Figure 1 depicts the Sioux Falls network with 24 nodes and 76 links.

After running the traffic assignment model on the Sioux Falls network, links and intersections (nodes) that have critical volume-capacity (v-c) ratios are identified as an initial set of project improvements. The BPR function (19) used as a link performance function allows v-c ratios to exceed 1.0, which helps us identify the most congested links.

The project alternatives considered are link widenings (which are assumed to be applied symmetrically in both directions between the two connecting nodes because the O-D table is symmetric), and vertical, horizontal, or vertical and horizontal, improvements of intersections. Improvements are carried through the entire intersection for consistency with the number of lanes on the intersection's legs; there are two types of improvements that are considered in this paper: (i) N-S widening of the intersection between the North–South approaches, (ii) E-W widening of the intersection between East–West approaches. It is assumed that some projects should be bundled because it saves costs due to the joint use of resources and construction equipment. The assumption to bundle some projects is justified economically because it saves costs due to the joint use of resources and construction equipment.

In this example, it is assumed that the demand grows exponentially over the planning horizon as

$$d_{ij}^t = d_{ij}^{0*}(1 + g)^t \qquad (9)$$

where $d^t{}_{ij}$ is the demand between origin i and destination j, $d^0{}_{ij}$ is the base demand for the ij origin and destination (O-D) pair at time 0, and g is the growth rate per period. Some numerical values of the input parameters and their units are displayed in Table 1.

Nodes 8, 11 and 16 represent two-phased intersections in the Sioux Falls network. Intersections were modeled by adding one pseudo-link for each movement between link pairs, for example, for intersection 8, link 47 there are three pseudo-links (47002, 47004, 47006) for three movements (left-turn (002 part), through movement (004), and right-turn (006), respectively). Overall, for the three intersections (8, 11 and 16), 36 pseudo-links are added to the network. Table 2 shows the pseudo-links for intersections, their capacity, free flow travel time (t0), and which pseudo-link belongs to which intersection. The capacity of each pseudo-link was set as the minimum value of the capacities of the two real links it connects.

Table 3 shows the initial volumes for each of the O-D pairs. It is evident that there are no trips originating and ending at nodes 8, 11, 16, because we consider them as intersections in the Sioux Falls network. In Table 4, the values of delay on intersection pseudo-links, the volumes on each pseudo-link, and pseudo v-c ratio are presented.

Figure 2 shows how the values of delay increase as the pseudo v-c ratio increases, for the three pseudo-links 36004, 16002, and 52002. These pseudo-links were chosen because of their large increases in delays as volume increases. The delay on each of the pseudo-links varies slightly, as can be seen in Figure 2.

Figure 3 shows overall intersection delay for the three intersections as function of the percentage of increase of the original O-D volumes, in 10% increments ranging from 10% to 140% of the original O-D table. In it the intersection delay usually increases as the percentage of volume increases, with intersections 8 and 16 having the greatest increases in delay. Due to traffic re-assignment, the delay increase is not monotonic at individual intersections.

Table 2. Pseudo-Links for Intersections 8, 11, and 16, Their Capacity and Free Flow Travel Time

1st Node	2nd node	Pseudo-link ID	Capacity	t_0	Intersection
4	14	10004	731	2.652	11
4	10	10002	736	3	
4	12	10006	736	3.876	
6	16	16004	734	1.302	8
6	9	16006	734	1.302	
6	7	16002	734	1.302	
7	9	17004	757	1.5	8
7	6	17006	734	1.302	
7	16	17002	756	1.5	
8	17	22004	756	1.002	8
8	10	22006	728	2.7	
8	18	22002	756	1.614	
12	4	36002	736	3.876	11
12	10	36004	736	3	
12	14	36006	731	2.652	
14	12	40002	731	2.652	11
14	4	40004	731	2.652	
14	10	40006	731	2.652	
16	9	47002	756	2.892	8
16	6	47004	734	1.302	
16	7	47006	756	1.5	
17	10	52002	728	1.002	16
17	8	52004	756	1.002	
17	18	52006	756	1.002	
18	17	55002	784	1.002	16
18	10	55004	728	1.614	
18	8	55006	756	1.614	

Intersections 8 and 16 are considered for improvement based on their delay values. The links to be improved were chosen because of their high v-c ratios (above 0.6). Table 5 summarizes the list of projects. Intersections with the highest delay values and links with the highest v-c ratio are selected for improvement. Table 6 shows the bottleneck sequence and schedule of projects (ordered based on the projects' v-c ratios), greedy sequence and schedule (ordered based on their benefit-cost (b-c) ratio), and the GA-optimized sequence and schedule of projects. In this case, benefit is defined as the monetary value of total travel time savings from implementing one project, and cost is simply the implementation cost of each project. The present values of total costs after each project implementation are also shown in Table 6. These results indicate that the GA yields a better solution, that is, with lower total cost compared to sequences based on b-c and v-c ratios. This occurs because the GA process accounts for project interrelations, unlike common practices such as b-c ratio and congestion level rankings.

Figure 4 shows the performance of the GA; the optimized solution is reached after 22 generations. The stopping criterion for the GA was set at 10 successive similar solutions (shown in Figure 4) but, for more confidence in the results, we let it run further for 200 generations, which yielded the same solution. The CPU time for entire analysis is 3300 seconds. Table 7 demonstrates the sensitivity of the optimized sequence, schedule, and the objective function value (total cost) to changes in demand.

Demand is changed by the same percentage for each cell in the O-D matrix. Table 7 also presents the sensitivity of results to changes in the available budget. The variation in budget is specified as different percentages of the original value, which was set to $1.5 million/year. It should be noted that unsteady budget flows do not increase the model's complexity or computation time.
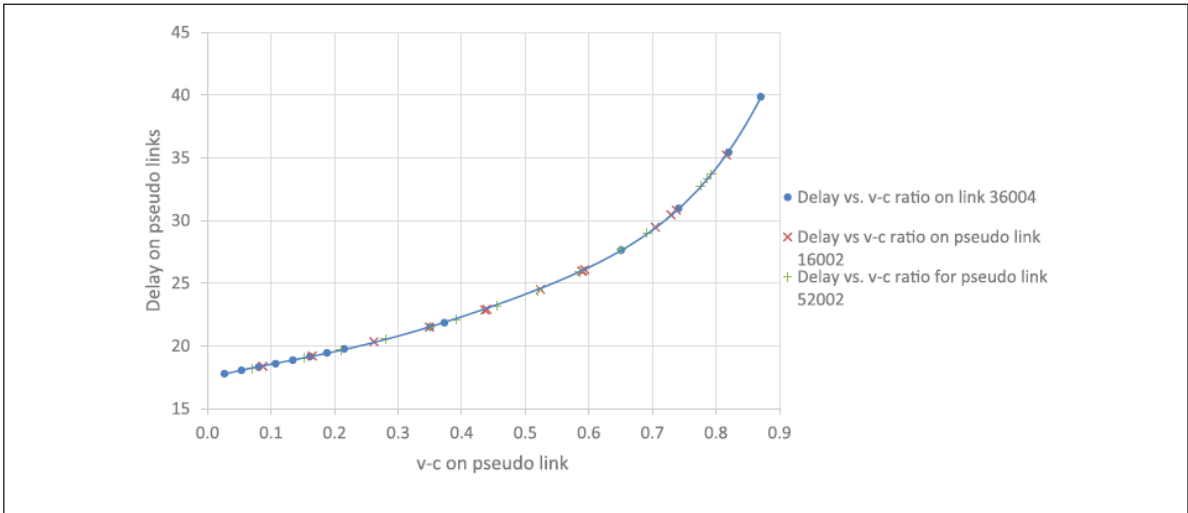
Figure 2. Delay vs. v-c ratio for pseudo-link 36004.

Table 3. Baseline O-D Volumes for the Sioux Falls Network

O-D	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	200	120	360	180	240	300	0	360	840	0	180	360	180	300	0	300	120	180	180	60	240	180	120
2	200	0	6	18	6	30	12	0	18	36	0	12	18	6	12	0	18	6	6	12	6	12	6	6
3	120	6	0	18	6	18	6	0	12	18	0	18	12	6	6	0	6	0	6	6	6	6	6	6
4	360	18	18	0	30	30	30	0	48	72	0	42	36	30	30	0	30	6	18	24	12	24	30	18
5	180	6	6	30	0	18	12	0	48	60	0	12	12	12	18	0	18	6	12	12	6	12	12	6
6	240	30	18	30	18	0	24	0	24	48	0	18	18	12	18	0	36	6	18	24	6	18	12	6
7	300	12	6	30	12	24	0	0	36	114	0	48	30	18	30	0	60	60	30	36	18	36	12	6
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	360	18	12	48	48	24	36	0	0	168	0	42	36	36	60	0	60	12	30	42	24	42	36	12
10	840	36	18	72	60	48	114	0	168	0	0	126	114	132	240	0	234	42	108	156	78	162	108	54
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	180	12	18	42	12	18	48	0	42	126	0	0	84	42	48	0	42	12	18	30	24	48	42	30
13	360	18	12	36	12	18	30	0	36	114	0	84	0	36	42	0	36	6	24	42	36	78	48	48
14	180	6	6	30	12	12	18	0	36	132	0	42	36	0	45	0	42	6	24	30	24	72	66	24
15	300	12	6	30	18	18	30	0	60	240	0	48	42	45	0	0	90	18	48	66	48	156	60	30
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	300	18	6	30	18	36	60	0	60	234	0	42	36	42	90	0	0	42	102	102	42	102	36	18
18	120	6	0	6	6	6	60	0	12	42	0	12	6	6	18	0	42	0	24	100	6	24	6	6
19	180	6	6	18	12	18	30	0	30	108	0	18	24	24	48	0	102	24	0	78	30	78	24	12
20	180	12	6	24	12	24	36	0	42	156	0	30	42	30	66	0	102	100	78	0	78	150	42	30
21	60	6	6	12	6	6	18	0	24	78	0	24	36	24	48	0	42	6	30	78	0	114	42	36
22	240	12	6	24	12	18	36	0	42	162	0	48	78	72	156	0	102	24	78	150	114	0	132	72
23	180	6	6	30	12	12	12	0	36	108	0	42	48	66	60	0	36	6	24	42	42	132	0	4.8
24	120	6	6	18	6	6	6	0	12	54	0	30	48	24	30	0	18	6	12	30	36	72	4.8	0

Table 4. Values of Delay [sec/veh], Volume [veh] and v-c Ratios for Pseudo-Links at Intersections 8, 11, and 16

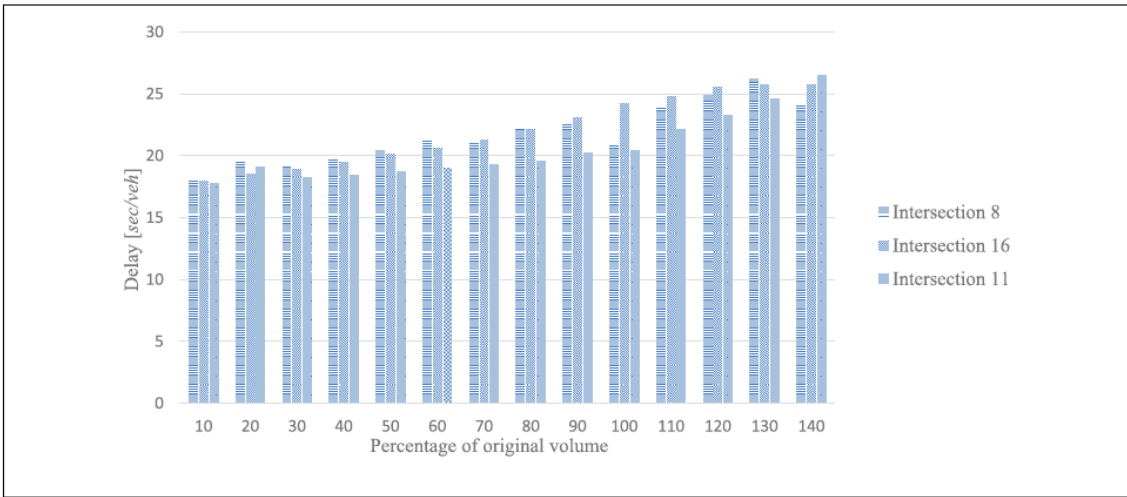| Intersection 8				Intersection 16				Intersection 11			
Pseudo-link	Delay	Volume	v-c	Pseudo-link	Delay	Volume	v-c	Pseudo-link	Delay	Volume	v-c
47002	17.97572	30.01127	0.039652	52002	27.89938905	479.8608	0.658934	40002	20.85631	222	0.303496
47004	21.64097	263.9938	0.359279	52004	19.50075055	142.3221	0.18804	40004	18.64281	83.99553	0.11483
47006	20.88573	228.6979	0.302161	52006	18.35787152	60.00031	0.079274	40006	19.53734	144	0.196862
24004	20.17247	186.009	0.245547	55002	20.00773625	178	0.2269	27002	21.30628	246	0.336306
24002	18.56341	78.01732	0.106177	55004	18.30268932	60.00649	0.0824	27004	21.94002	279.2566	0.379258
24006	17.97555	29.99829	0.039634	55006	19.52566035	144	0.190256	27006	17.542	0	0
17004	18.09034	108.0194	0.142594	22004	19.33305523	130.9015	0.17295	10004	18.52118	75.35613	0.103019
17006	19.00682	443.984	0.604234	22006	17.51455175	0	0	10002	17.542	0	0
17002	26.32533	203.9935	0.269521	22002	19.61527684	149.9971	0.19818	10006	19.91571	168.0003	0.228161
16004	20.46652	78.98595	0.107495	29002	18.08119079	43.67763	0.059977	36002	19.62855	150.0013	0.203716
16006	18.3156	60.00376	0.081661	29004	18.30256332	59.99731	0.082387	36004	22.87071	322.163	0.437529
16002	25.95317	433.1572	0.589499	29006	29.63127316	515.9972	0.708556	36006	20.85642	222.0059	0.303504

Figure 3. Intersection delay vs. percentage of the original volume of the O-D table.

Table 5. List of Candidate Projects and Their Costs

Project ID	Project description	Project cost
1	Improvement of links 69 & 65	$1,800,000.00
2	Improvement of links 30 & 51	$4,800,000.00
3	Improvement of links 62 & 64	$3,900,000.00
4	Improvement of links 68 & 63	$4,200,000.00
5	horizontal improvement of intersection 8 (pseudo-links 17, 24)	$220,880.00
6	vertical improvement of intersection 8 (pseudo-links 16, 47)	$220,880.00
7	horizontal improvement of intersection 16 (pseudo-links 29, 55)	$220,880.00
8	vertical improvement of intersection 16 (pseudo-links 22, 52)	$220,880.00

Table 6. Bottleneck, Greedy, and GA-Optimized Schedule of Projects with Corresponding Total Costs

| Bottle neck sequence	Present value of total cost ($)	Schedule (years)	Greedy order sequence	Present value of total cost ($)	Schedule (years)	GA- optimized sequence	Schedule (years)	Present value of total cost ($)
1	532,870,256	1.2	2	1,295,746,013	3.2	7	0.15	68,455,247
2	1,701,652,369	4.4	3	2,095,244,334	5.8	6	0.29	135,967,546
3	2,421,799,340	7	6	2,134,903,906	5.95	2	3.49	1,396,017,082
4	3,025,641,095	9.8	7	2,174,285,216	6.09	3	6.09	2,174,056,728
5	3,053,349,907	9.95	8	2,213,386,599	6.24	4	8.89	2,825,298,019
6	3,080,981,763	10.09	5	2,252,222,345	6.39	1	10.09	3,057,265,698
7	3,108,535,464	10.24	4	2,902,828,902	9.19	5	10.24	3,085,002,834
8	3,136,038,746	10.39	1	3,134,797,942	10.39	8	10.39	3,112,075,181
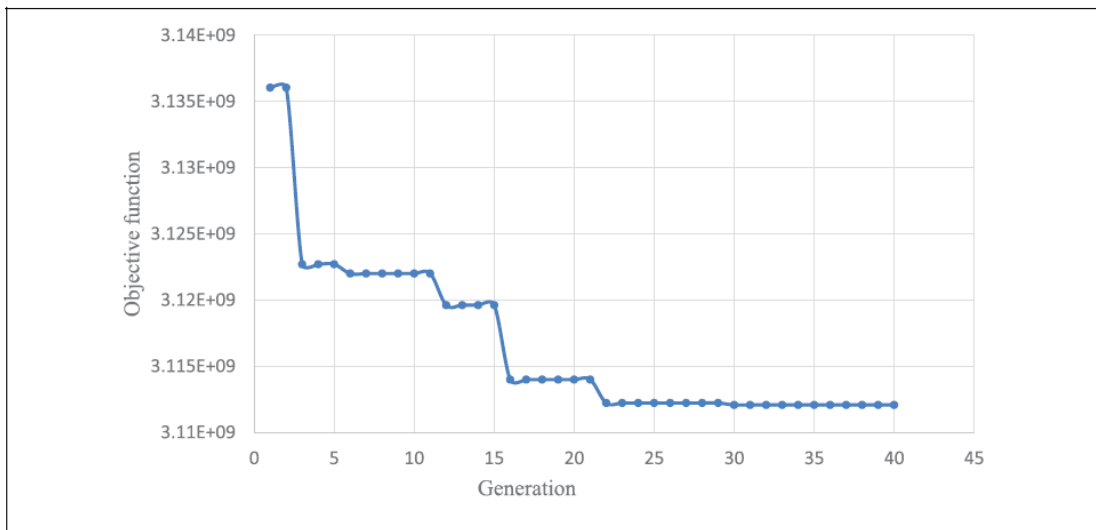
Figure 4. Performance of the genetic algorithm.

38

Conclusion

The improvement of intersections and links in a network is just one example of interrelated alternatives for which the selections and scheduling of projects becomes a challenging optimization problem. This paper modifies the FW traffic assignment model to consider intersection flows and delays. This is done by introducing pseudolinks to the network and applying Akcelik's delay model. The modified model is then incorporated within a GA loop to optimize the selection and scheduling problem. Common prioritizing practices which are rankings based on b-c ratio and congestion level do not produce the optimal sequence of projects because they disregard the interrelations among projects, unlike the GA used here. This methodology can be applied more generally to other more complex cases. GAs can optimize very intractable objective functions without requiring restrictive assumptions about their structures which allows them to be efficiently combined with other evaluation tools, to solve selecting and scheduling problems.

Future research may focus on extending the model by incorporating more detailed evaluation methods (such as simulation models) to capture dynamic effects in congested networks that are missed by the FW algorithm. Future model versions may also consider more elaborate intersection configurations, control policies and cyclical variations in daily and weekly traffic. Gas could be solved considerably faster by distributing the evaluation of population members among multiple processors. Moreover, individual improvements (resurfacing, widening) could be grouped to form a project, bus traffic could be traced along with passenger vehicles in the traffic assignment method, and different cost rates could be assumed for different types of improvements implemented.

Table 7. Sequence and Schedule Change Subject to Demand Change and Sequence and Schedule Change Subject to Budget Change

	Sequence	Schedule (years)	Present value of total cost
Demand			
50%	6-7-3-2-4-1-8-5	0.1-0.3-2.9-6.1-8.9-10.1-10.2-10.4	$1,532,306,913
75%	7-6-2-4-3-1-5-8	0.1-0.3-3.5-6.3-8.9-10.1-10.2-10.4	$2,309,379,871
100%	7-6-2-3-4-1-5-8	0.1-0.3-3.5-6.1-8.9-10.1-10.2-10.4	$3,112,075,181
110%	7-6-2-4-3-1-8-5	0.1-0.3-3.5-6.3-8.9-10.1-10.2-10.4	$3,452,620,779
125%	7-6-2-3-4-1-8-5	0.1-0.3-3.5-6.1-8.9-10.1-10.2-10.4	$3,989,723,752
Budget			
50%	6-7-2-4-3-1-8-5	0.3-0.6-7-12.6-17.8-20.2-20.5-20.8	$4,272,866,372
66%	6-7-2-4-3-1-5-8	0.2-0.4-5.2-9.4-13.3-15.1-15.4-15.6	$3,832,339,985
100%	7-6-2-3-4-1-5-8	0.1-0.3-3.5-6.1-8.9-10.1-10.2-10.4	$3,112,075,181
133%	7-6-3-2-4-1-5-8	0.1-0.2-2.2-4.6-6.7-7.6-7.7-7.8	$2,594,955,256
200%	7-6-3-2-4-1-5-8	0.07-0.1-1.4-3-4.4-5-5.1-5.2	$1,935,240,006

Acknowledgments

References

1. Shayanfar, E., A. S. Abianeh, P. Schonfeld, and L. Zhang. Prioritizing Interrelated Road Projects Using Metaheuristics. Journal of Infrastructure Systems, Vol. 22, No. 2, 2016, p. 04016004.

2. Wardrop, J. G. Road Paper. Some Theoretical Aspects of Road Traffic Research. Proceedings of the Institution of Civil Engineers, Vol. 1, No. 3, 1952, pp. 325–362. 3. Webster, F. V. Traffic Signal Settings. Road Research Technical Paper No. 39. Road Research Laboratory, London, 1958.

4. Highway Capacity Manual. TRB, National Research Council, Washington, D.C., 2000.

5. Akcelik, R. The Highway Capacity Manual Delay Formula for Signalized Intersections. ITE Journal, Vol. 58, No. 3, 1988, pp. 23–27.

6. Akcelik, R. Appendix: A Note on the Generalized Delay Model. Compendium of Technical Papers, 60th Annual Meeting of Institute of Transportation Engineers, Orlando, Fla., 1990, pp. 29–32.

7. Teply, S. Quality of Service in the New Canadian Signal Capacity Guide. Proc., International Symposium on Highway Capacity and Level of Service, A. A. Balkema, Rotterdam, 1991, pp. 377–386.

8. Vincent, R.A., Mitchell, A.I., Roberston, D.I. User Guide to TRANSYT Version 8. Transportation and Road Research Laboratory (TRRL), Workingham, Berkshire United Kingom, 1980.

9. Heidemann, D. Queue Length and Delay Distributions at Traffic Signals. Transportation Research Part B: Methodological, Vol. 28, No. 5, 1994, pp. 377–389. https://doi.org/ 10.1016/0191-2615(94) 90036-1

10. Olszewski, P. S. Modeling Probability Distribution of Delay at Signalized Intersections. Journal of Advanced Transportation, Vol. 28, No. 3, 1994, pp. 253–274. https://onlinelibrary.wiley.com/doi/abs/10.1002/atr.5670280306

11. Weingartner, H. M. Capital Budgeting of Interrelated Projects: Survey and Synthesis. Management Science, Vol. 12,No. 7, 1966, pp. 485–516. https://doi.org/10.1287/mnsc.12.7.485

12. Cochran, M. A., E. B. Pyle, L. C. Greene, H. A. Clymer, and A. D. Bender. Investment Model for R&D Project Evaluation and Selection. IEEE Transactions on Engineering Management, Vol. EM-18, No. 3, 1971, pp. 89–100. https://dx.doi.org/10.1109/TEM.1971.6447136

13. Nemhauser, G. L., and Z. Ullmann. Discrete Dynamic Programming and Capital Allocation. Management

18. Ran, B., and D. Boyce. Modelling Dynamic Transportation Networks: An Intelligent ransportation System Oriented Approach. Springer-Verlag, Berlin, Heidelberg, 1996.

19. Frank, M., and P. Wolfe. An Algorithm for Quadratic Programming. Naval Research Logistics (NRL), Vol. 3, No. 1–2, 1956, pp. 95–110. http://dx.doi.org/10.1002/nav .3800030109

20. Van Leeuwen, J. Handbook of Theoretical Computer Science (Vol. A): Algorithms and Complexity. The MIT Press/Elsevier, Cambridge MA/Amsterdam, The Netherlands, 1991.

21. Dorigo, M., and T. Sttzle Ant Colony Optimization. The MIT Press, Cambridge, Mass., 2004.

22. Jong, J. C., and P. Schonfeld. Genetic Algorithm for Selecting and Scheduling Interdependent Projects. Journal of Waterway, Port, Coastal, and Ocean Engineering, Vol. 127, No. 1, 2001, pp. 45–52.

23. Teodorovic, D. Transportation Networks. University of Belgrade, Faculty of Transport and Traffic Engineering, Belgrade, 2007.

24. LeBlanc, L. J., E. K. Morlok, and W. P. Pierskalla. An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem. Transportation Research, Vol. 9, No. 5, 1975, pp. 309–318. https://doi.org/10.1016/0041-1647(75)90030-1

Optimizing development plan of rail transit projects over multiple time periods

Ya-Ting Peng[a,b], Zhi-Chun Li[a], Paul Schonfeld[b]

[a]*School of Management, Huazhong University of Science and Technology, Wuhan 430074, China*

[b]*Department of Civil and Environmental Engineering, University of Maryland, 1173 Glenn Martin Hall, College Park, MD 20742, USA*

Abstract

This paper addresses the development of interrelated rail transit projects in urban rail transit networks over multiple time periods. It extends the traditional network design problems by explicitly considering the time horizon and interrelations among projects in rail transit networks. The proposed model determines which projects in a rail transit network should be selected and completed at what times (i.e., project selection, sequence and completion time), while jointly optimizing the evolving headways of rail transit lines, in order to minimize the present value of the total cost. In addition to the financial budget provided by relevant agencies (e.g., governments), we consider fare revenues generated from the operations of previous completed projects as an internal source of funding for later projects. A Genetic Algorithm (GA) is adapted to solve this model and tested on the transit network development of Wuhan city in China. Sensitivity analysis is conducted to explore the effects on the development plan of some important factors, such as travel demand and annual financial budget. Findings are reported on the efficiency of the adapted GA approach as well as on the impacts of travel demand and budgets.

1. Introduction

The past decade has witnessed rapid growth in rail transit investments in China. According to the latest report by the Chinese Urban Rail Transit Association (CURTA, 2017), by the end of 2017, 165 rail transit lines with a total length of 5033 kilometers were operating in 34 cities in mainland China. Currently, 5636 km of rail transit lines are under construction, and 7305 km of rail lines were approved but not yet built. These rail transit projects require huge investment costs. For example, the capital cost of Wuhan Metro Line 2 was about RMB600 million per kilometer (RMB is the Chinese currency "Renminbi". US$1 approximates RMB6.51 as of January 1, 2018). However, the government funds available for investment in rail transit projects are limited. The investment or improvement of the rail transit lines is thus usually a multi-stage process.

As an example, Fig. 1 shows the gradual development process of the rail transit projects in Wuhan (a city located in Central China) in the past dozen years. It can be seen that Wuhan's rail transit network gradually expands from one line in 2005 to seven lines in 2017. The corresponding total rail line length grows from 34.57 km to 237 km. During the development process, the order and time of the project implementations can significantly affect user cost and the investment efficiency in terms of total cost. This raises an important question addressed here: how should we design an appropriate development plan for rail transit projects within financial constraints over a planning horizon such that the discounted total cost in the urban system is minimized?

In the literature, transportation infrastructure investment issues have attracted widespread interest due to their practical importance. Table 1 summarizes some principal contributions to the related problems, in terms of the type of infrastructure, consideration of time horizon, and consideration of interrelations among projects. It can be seen from Table 1 that the existing studies mainly focused on the general road network design problems with a discrete approach (see e.g., Wang et al., 2013; Zhang et al., 2014; Wang et al., 2015), a continuous approach (see e.g., Li et al., 2012; Yin et al., 2014; Liu and Wang, 2015), or in a hybrid way (see e.g., Luathep et al., 2011). These models usually aimed to add new links or expand the capacities of the old ones in the network. Certainly, this is also an important part of urban rail transit network development. However, the urban rail transit network development problem is more complex than the general road network development problems due to the design and operating characteristics of rail transit lines. In this regard, Gao et al. (2004) developed a bi-level model to examine the interaction between the supply side and the demand side in a transit network design problem. Farahani et al. (2013) provided a comprehensive review of urban transportation network design problems.

However, most of these were static models focused on stationary states, which cannot address the dynamic or progressive improvements of the rail transit system. It is well known that as the urban economy and population grows, together with the development for the transit network, the demand for the rail transit service may significantly increase. This increase can affect the rail services such as their headways, operating costs and fare revenues. Hence, the development decisions for the rail transit network should change, which in turn affect the system's travel demand. Thus, the demand for rail transit service, the operational condition and the network development decisions in one period are significantly affected by the decisions made in the previous periods, and therefore, vary over the entire time horizon. Consequently, it is important to incorporate the time dimension in the rail transit network development problem such that interactions between the supply and demand over different time periods can be taken into account.

So far, researchers have made considerable efforts to consider the time horizon in transport network design problems. For example, Cheng and Schonfeld (2015) optimized the extension of single rail line outward from a city center over time. Shayanfar et al. (2016) proposed an optimization framework for selecting and scheduling interrelated projects in a road network. Sun et al. (2017) explored the selection of public transit modes by costs and benefits analysis and considered essential factors in a long-term planning process, such as economies of scale in rail extensions and future cost discounting. More recently, Sun et al. (2018) extended the work of Cheng and Schonfeld (2015) by developing a bi-level model to determine how many stations along a rail line should be completed in different time periods, while considering demand elasticity. It should be noted that the previous relevant studies only considered single rail line, expanded outward from a city center. No comparable studies have been found for the more general rail transit network development problem.

In this paper, we extend the related studies to consider the gradual development process of urban rail transit networks, while accounting for correlations among projects in the rail transit network over different time periods. Here, a project means to invest in one segment or link in a rail transit network. Correlations among projects occur when the benefits and costs of projects in the rail transit network depend on whether and when other projects are completed. When a project is implemented, both the user costs of the newly built segments and those of the completed segments change since the number of OD pairs connected by rail lines and thus the demand for rail services increases. Growing travel demand can decrease the train headways and thus the user costs of completed segments along the rail lines. However, the operating costs increase due to the rail transit network expansion and decreasing train headways. Consequently, the total cost change (or project benefit) due to project development is not a simply linear summation of cost changes from individual segments, but a consideration of the operating cost increases and the user cost savings from all segments in the network. The correlations among projects significantly affect the investment decision and the development plan. Thus, it is important to account for the correlations among projects in the transit network and their effects on the system's total cost.

In light of the above discussion, this paper proposes a model for optimizing transit network development process over time by considering time-varying demand, financial constraints, and interrelations among projects over time and space. There are two main contributions in this paper. First, a novel model is proposed to determine the development process of rail projects in a rail transit network with limited financial budget over a planning horizon. In the proposed model, the present value of the total cost is minimized by optimizing the project selection, sequence and implementation schedule. The effects of the newly completed projects on transit systems and the present value of the total cost are explicitly explored by incorporating the correlations among projects over time and space. In addition, the growth of the travel demand over time is effectively captured by a time-varying travel demand function. The budget constraint includes possible internal funding, such as from the fare revenue generated from the operation of the transit rail lines. In other words, in addition to externally provided budgets, the fare revenue collected from the previous years is used as an internal source of funding to finance the successive projects. Second, some important factors that affect the development plan of the public transit projects and the present value of the total system cost are identified. Results reveal that both the initial travel demand and annual financial budget can significantly affect the development plan for a rail transit network. The proposed model can serve as a useful tool to guide the development process of urban transit networks.

The remainder of this paper is organized as follows. The next section describes some basic assumptions and the components of the models, including user cost and supplier cost. Section 3 presents the model for optimizing the development plan by determining which projects will be selected, when these projects are completed, and the train headways in each period on the rail lines in the network. A genetic algorithm (GA) for solving the proposed model is presented in Section 4. Next, numerical examples are provided to illustrate the applications of the proposed model in Section 5. Finally, Section 6 provides conclusions and recommendations for further studies.

2. Components of the model

2.1. Assumptions

To facilitate the presentation of essential ideas without loss of generality, some basic assumptions are made as follows.

A1. The layouts of rail transit lines and station locations are assumed to be exogenously given, as assumed in Cheng and Schonfeld (2015) and Sun et al. (2018). In fact, determining the layouts of rail transit lines and station locations in an urban rail transit network is a major task of transit system planning. In this paper, we focus on the future development plan for this pre-given transit network, that is, determining which projects should be selected and when these projects should be invested over a planning horizon.

A2. It is assumed that the sequenced projects can be invested once the financial budget is available. We aim to explore the transit network development by considering financial feasibility over time. Moreover, the system operations such as rail line length and train headways change if new projects are completed. These assumptions have been adopted in various previous studies (see e.g., Wang and Schonfeld, 2008; Shayanfar et al., 2016).

A3. Travel demand is assumed to be at a stationary state within each development period but varies among periods. Here, period refers to the development state of a transit network. Specifically, when a project is completed (i.e., the development state of the network changes), the current period ends and the next period begins. Therefore, the duration of periods depends on the interval between the completion of two successive projects, which is determined by the development plan and may vary over different periods. It is assumed that travel demand in different periods increases due to demographic trends, economic growth and network development. It is also assumed that the travel demand between OD pairs which are already connected by rail lines increases at a higher rate than that between unconnected OD pairs. In this paper, an exponential form of travel demand function is adopted (as in e.g., Shayanfar et al., 2016; Cheng and Schonfeld, 2015; Sun et al., 2018).

A4. The present value of the total cost in the urban system is assumed to be the sum of the discounted total cost over all development periods (see e.g., Shayanfar et al., 2016). In each period, the total cost includes user cost and supplier cost. The supplier cost refers to the cost for providing transit service, which includes the capital investment, network maintenance, and vehicle operating cost.

A5. It is assumed that until origin-destination pairs are connected by rail lines, their demands are served by other modes (e.g. autos or buses), at a cost proportional to travel distance.

A6. In this model at most one rail route exists between any OD pair. In fact, except in central parts of cities with very large rail networks, most rail trips have no alternative rail paths. This typical situation can be seen in many cities, such as Atlanta.

2.2. User cost

Consider an urban rail transit network $G(N, A)$, where N is the set of nodes (transit stations or stops) and A is the set of transit line segments in the network. Let W be the set of origin-destination (OD) pairs in the network, L be the set of transit lines and T be the set of development periods. The binary decision variable can be defined as

$$y_a^{(t)} = \begin{cases} 1, \text{if segment } a \text{ already exists in period } t, a \in A, t \in T, \\ 0, \text{otherwise.} \end{cases} \tag{1}$$

It should be noted that in the rail transit network, segment a may include several stations. This is consistent with actual practice because it can yield economies of scale and save costs in using mobilized resources such as construction equipment.

Let $c_{a1}^{(t)}$ and $c_{a2}^{(t)}$ be the user cost on segment a by rail and by other modes in period t, respectively. The travel cost by rail consists of waiting cost and in-vehicle time cost. Note that the access cost that be omitted because we assume that the station locations are predetermined (see Assumption 1). Thus, we have

$$c_{a1}^{(t)} = \lambda_1 \frac{d_a}{V} + \lambda_2 \frac{\chi_{al} h_l^{(t)}}{2}, a \in A, l \in L, t \in T, \tag{2}$$

where λ_1 and λ_2 are the values of in-vehicle time and waiting time, respectively. χ_{al} is a 0-1 indicator, which equals 1 when segment a is a section of rail line l, and 0 otherwise. d_a is the length of segment a, V is the average speed of trains, and $h_l^{(t)}$ is the average train headway of rail line l where segment a is located in period t. According to Assumption 5, the user cost on segment a by other modes, $c_{a2}^{(t)}$, can be expressed as

$$c_{a2}^{(t)} = c_0 d_a, a \in A, t \in T, \tag{3}$$

where c_0 is the cost per km of travelling by other modes, which is assumed to be a constant. It can be seen from Assumption 5 that, before segment a is implemented or connected to rail lines, persons passing through it have to choose other travel modes. Let $c_a^{(t)}$ be the user cost on segment a, which can be expressed as

$$c_a^{(t)} = \left(1 - y_a^{(t)}\right) c_{a1}^{(t)} + y_a^{(t)} c_{a2}^{(t)}, a \in A, t \in T, \tag{4}$$

where $y_a^{(t)}$ is the decision variable, defined in Eq. (1), indicating whether segment a is completed in period t.

The daily traffic volume on segment a in period t, $Q_a^{(t)}$, can be expressed as

$$Q_a^{(t)} = \sum_{w \in W} q_w^{(t)} \delta_{wa}^{(t)}, a \in A, w \in W, t \in T, \tag{5}$$

where $q_w^{(t)}$ is the daily travel demand between OD pair w in period t. $\delta_{wa}^{(t)}$ is an indicator, which equals 1 when segment a is on the route between OD pair w in period t, and 0 otherwise. Note that there is at most one rail route connecting OD pair w (see Assumption 6). Therefore, the route index is omitted here. We assume the travel demand increases over time due to demographic and economic growth and network development. According to Assumption 3, the exponential form of travel demand function can be expressed as

$$q_w^{(t)} = q_w^{(0)} \left(1 + g_1\right)^{x_t} (1 + g_2 \zeta_w)^{x_t - x_w}, t \in T, w \in W, \tag{6}$$

where $q_w^{(0)}$ is the daily travel demand between OD pair w in period 0, g_1 is the base growth rate per year due to demographic and economic growth and g_2 is the additional annual growth rate when OD pair w is connected (see Assumption 3). ζ_w is a 0-1 indicator, which equals 1 when OD pair w is connected, and 0 otherwise. x_t is the starting time of period t, and x_w is the first time to complete the connection for OD pair w. Let $C_u^{(t)}$ be the annual user travel cost in period t. Thus, we obtain

$$C_u^{(t)} = \rho \sum_{a \in A} Q_a^{(t)} c_a^{(t)}, a \in A, t \in T, \tag{7}$$

where ρ is the average number of days of travel per traveler per year, which is used to transform the daily basis cost to the yearly one. $Q_a^{(t)}$ is the daily traffic volume on segment a in period t and $c_a^{(t)}$ is the user cost on segment a.

2.3. Supplier cost

According to Assumption 4, the cost of providing the rail transit service in each period includes the capital investment cost of the new project, the maintenance cost of existing rail lines in this period, and the vehicle operating cost in this period. Let $\Lambda_c^{(t)}$ be the capital investment cost in period t, $\Lambda_m^{(t)}$ be the annual maintenance cost in period t, and $\Lambda_o^{(t)}$ be the annual vehicle operating cost in period t. The capital investment cost $\Lambda_c^{(t)}$ in period t, such as land acquisition, design, and construction costs, can be expressed as

$$\Lambda_c^{(t)} = \sum_{a \in A} \left(y_a^{(t+1)} - y_a^{(t)}\right) \phi_a, t \in T, \tag{8}$$

where ϕ_a is the capital investment cost for segment a. Note that the capital investment cost only occurs at the time when segment a is developed (i.e., the end of this period and the beginning of the

47

next period). Here, the term $\left(y_a^{(t+1)} - y_a^{(t)} \right)$ indicates whether or not segment a is selected at the end of period t. It equals 1 when segment a is implemented in period t, and 0 otherwise.

The maintenance cost $\Lambda_m^{(t)}$ per year in period t, is directly proportional to the total length of the existing transit lines in period t, which can be expressed as

$$\Lambda_m^{(t)} = \eta \sum_{a \in A} y_a^{(t)} d_a, \tag{9}$$

where η is maintenance cost of transit lines per kilometer per year.

The annual vehicle operating cost is the sum of the vehicle operating cost of each transit line. Specifically, the annual vehicle operating cost of a transit line is its fleet size multiplied by annual operating cost per train. To obtain the fleet size, the transit round trip time should be derived first. Let $R_l^{(t)}$ be the round trip time of line l in period t and $F_l^{(t)}$ be the fleet size of transit line l in period t. Thus,

$$R_l^{(t)} = 2 \sum_{a \in A} \left(d_a y_a^{(t)} \chi_{al} \right) \Big/ V, \, l \in L, t \in T, \tag{10}$$

$$F_l^{(t)} = R_l^{(t)} \Big/ h_l^{(t)}, \, l \in L, t \in T, \tag{11}$$

where χ_{al} is a 0-1 indicator determining whether or not segment a is a section of rail line l, defined in Eq. (2). $\sum_{a \in A} \left(d_a y_a^{(t)} \chi_{al} \right)$ is the length of line l completed in period t, which may change due to the network development. Let β be the operating cost per train per year. Therefore, the total yearly vehicle operating cost of the system in period t $\Lambda_o^{(t)}$ can be expressed as

$$\Lambda_o^{(t)} = \sum_{l \in L} \beta F_l^{(t)}, t \in T. \tag{12}$$

The rail line's headway varies with its travel demand. Consequently, the headways are steady in each development period, but vary among periods, like the changes in travel demand (see Assumption 3). Therefore, we have to re-optimize the headways in each period, i.e. after every decision made. The optimal headway for rail line l in period t $h_l^{(t)}$, can be determined by minimizing the total cost of the system in this period. Specifically, the system's total cost in period t is defined as the sum of the user cost and the supplier cost in this period. Let $\Omega^{(t)}$ be the total cost of the system in period t. According to Eqs. (4)-(12), it can be expressed as

$$\Omega^{(t)} = \Delta_t \left[\rho \sum_{a \in A} Q_a^{(t)} \left[\left(1 - y_a^{(t)} \right) c_{a1}^{(t)} + y_a^{(t)} c_{a2}^{(t)} \right] + \eta \sum_{a \in A} y_a^{(t)} d_a + \sum_{l \in L} \beta R_l^{(t)} \Big/ h_l^{(t)} \right] + \sum_{a \in A} \left(y_a^{(t+1)} - y_a^{(t)} \right) \phi_a, t \in T, \tag{13}$$

$$\Delta_t = x_{t+1} - x_t, t \in T, \tag{14}$$

where Δ_t is the duration of period t, which is determined by the difference between the start time x_{t+1} of period $t+1$ and the start time x_t of period t. In square brackets in the right hand side of Eq. (13) the

first term represents the annual user cost in period t, the second term is the annual maintenance cost in period t, and the third term is the annual operating cost in period t. Setting $\dfrac{\partial \Omega^{(t)}}{\partial h_l^{(t)}} = 0$, we can analytically obtain the optimal headway of transit line l in period t as

$$h_l^{(t)} = \sqrt{\dfrac{2\beta R_l^{(t)}}{\lambda_2 \rho \sum\limits_{a \in A}\left(y_a^{(t)} Q_a^{(t)} \chi_{al}\right)}}, l \in L, t \in T, \tag{15}$$

where β is the operating cost per train per year and λ_2 is the value of waiting time. $\sum\limits_{a \in A}\left(y_a^{(t)} Q_a^{(t)} \chi_{al}\right)$ is traffic volume of line l in period t. Eq. (15) implies that the optimal headway of transit line l in period t, $h_l^{(t)}$, decreases to accommodate the increased demand of this line over time.

3. Model formulation

As previously stated, the goal is to minimize the present value of the total cost by determining which projects should be developed and when these projects should be completed. The discounted total cost is the sum of the discounted total cost in each period. According to Eqs. (4)-(13), the model can be formulated as follows.

$$\min_{\left(y_a^{(t)}, x_t\right)} \Theta = \sum_{t \in T} \dfrac{\Omega^{(t)}}{\left(1+r\right)^{x_t}} = \sum_{t \in T} \dfrac{\Delta_t \left(\rho \sum\limits_{a \in A} Q_a^{(t)} c_a^{(t)} + \eta \sum\limits_{a \in A} y_a^{(t)} d_a + \sum\limits_{l \in L} \beta R_l^{(t)} / h_l^{(t)}\right) + \sum\limits_{a \in A}\left(y_a^{(t+1)} - y_a^{(t)}\right)\phi_a}{\left(1+r\right)^{x_t}}, \tag{16}$$

s.t.

$$z_i^{(t)} + z_j^{(t)} \geq y_a^{(t+1)}, i, j \in N, a \in A, t=0,1,2,...,T-1, \tag{17}$$

$$z_i^{(t)} \geq y_a^{(t)}, t \in T, a \in A, i \in N, \tag{18}$$

$$z_j^{(t)} \geq y_a^{(t)}, t \in T, a \in A, j \in N, \tag{19}$$

$$z_n^{(t+1)} \geq z_n^{(t)}, n \in N, t=0,1,2,...,T-1, \tag{20}$$

$$y_a^{(t+1)} \geq y_a^{(t)}, a \in A, t=0,1,2,...,T-1, \tag{21}$$

$$\tau \sum_{a \in A}\left(y_a^{(t)} Q_a^{(t)} \chi_{al}\right) \leq \dfrac{K_{veh}}{h_l^{(t)}}, a \in A, t \in T, l \in L, \tag{22}$$

$$B^{(t)} + \Phi^{(t)} \geq \Lambda_c^{(t)}, t = 0,1,2,...,T, \tag{23}$$

$$B^{(t)} = B_0\left(x_t - x_{t-1}\right), t = 1,2,3,...,T, \tag{24}$$

where r is the discount rate. The denominator $(1+r)$ in Eq. (16) is used to convert the cost of future investment to today's cost. $y_a^{(t)}$ and x_t are the decision variables defined in Eqs. (1) and (6), respectively. Eq. (16) is the objective function that minimizes the present value of the system's total cost. $\mathbf{z}^{(t)} = \left(z_n^{(t)}, n = 1, 2, ..., N \right)$ is the vector of 0-1 variables indicating whether a node is completed in period t. i and j denote the indices for the two end nodes of segment a. Constraint (17) expresses the segment connectivity in the network, which implies that the segments to be built should have at least one end node already completed (i.e., the newly built segments must connect to the segments that have been already completed). This constraint ensures that the network's rail lines are extended by connecting to the existing lines. However, there is an exception. Initially, when none of nodes or segments in the network are yet completed, i.e., no existing lines need to be connected, any projects may be considered for immediate implementation without subject to Constraint (17). Constraints (18)-(19) mean that if segment a is completed, its two end nodes are also completed. Constraints (20) and (21) are realistic constraints ensuring that after nodes and segments are completed, they always remain in service in later periods. τ is the peak-hour factor, i.e., the ratio of peak-hour demand to the daily demand, which is used to convert the passenger volume from a daily basis to an hourly basis. K_{veh} is the capacity of vehicles (i.e., the maximum number of passengers allowed in a vehicle, both seated and standing). Constraint (22) is the line capacity constraint, which guarantees that the rail service supply satisfies the associated (peak-hour) passenger demand. $B^{(t)}$ is the budget flow in period t and B_0 is the annual budget level provided by relevant agencies (e.g., governments). Constraint (23) is a reformulated budget constraint which considers an internal funding source, such as the rail fare revenue collected from the rail service operations. The left-hand side of Constraint (23) denotes the total available funding at the end of period t and the right-hand side denotes the capital investment cost needed. The reformulated budget constraint reflects interrelations among projects in the transit network since the capital used for development is partly supplied by fare revenue collected from the rail operations, which may change with the network development. $\Phi^{(t)}$ denotes the fare revenue collected from the rail operations in period t, which can be expressed as

$$\Phi^{(t)} = \Delta_t \left[\rho \sum_{a \in A} \left(y_a^{(t)} Q_a^{(t)} f d_a \right) \right], \tag{25}$$

where Δ_t is the duration time of period t, defined in Eq. (14). The fare on segment a is the fare per km f multiplied by its length d_a.

It should be noted that if the budget is limited throughout the planning horizon, i.e., never sufficient for all beneficial projects, a project sequence uniquely determines a project schedule. The available funds should always be used whenever they suffice to complete a project (see Assumption 2). Hence, after the sequence of projects is determined, the completion time of these projects can be obtained by checking budget constraint. Accordingly, only those projects whose implementation times are within the planning horizon are selected. Here, the projects that are completed at the time beyond the planning horizon are implicitly rejected. Thus, the development plan is optimized by first optimizing the sequence of projects, and then determining the completion time of each project.

4. Solution algorithm

The above total cost minimization model (16)-(24) is a constrained integer programming problem, which is non-linear and non-convex, making it difficult to find its globally optimal solution. A GA approach is presented in this section due to its suitability for very "noisy" objective functions. GA's are inspired by phenomena in evolutionary biology. In a GA, a solution of the problem is called an individual. It is represented as a sequence of variables called a chromosome or gene string. A group including multiple individuals is defined as population. The essence of GA is population evolution through selection, crossover and mutation. Generally, a GA starts from initializing a set of individuals, i.e., a population, and then selecting the better individuals to reproduce offspring by applying genetic operators such as crossover and mutation operators. As a result, the most adapted individuals survive and thus the population can converge toward an optimized solution.

The GA in this paper is developed from basic GAs but differs from them in many ways. First, an efficient genetic encoding scheme is adopted to deal with the constraints. Since the proposed model has the network connectivity constraint (see Eq. (17)), traditional representation schemes such as the sequence of projects may generate too many infeasible solutions. A general remedy for this problem is to add penalty terms to fitness functions or use repair operators to transform infeasible solutions into feasible ones. However, these methods cannot handle the connectivity constraint efficiently and degrade the search efficiency in terms of speed and accuracy. Therefore, a novel genetic encoding scheme is needed. Second, solutions capturing the characteristics of the network and projects are incorporated into the initial population to accelerate the convergence of the GA. For example, solutions that represent the sequence of projects ordered by their demand level and investment cost are included in the initial population. Intuitively, development of projects with higher travel demand and lower investment cost can contribute more to the system cost saving and thus those projects have higher priority for development. As a result, such solutions may make better use of existing information, which help accelerate the convergence of the GA. Third, some mechanisms are designed to avoid GA prematurity. In the selection process, a ranking method is used to help the GA escape from local optima. In addition, the catastrophe mechanism is introduced when the optima remain unchanged for a certain number of generations (e.g., 50 generations). These mechanisms are capable of enhancing the accuracy and stability of the GA.

4.1. Genetic encoding and decoding

The process of encoding a chromosome into a string is called genetic encoding and the process of decoding a chromosome into a feasible solution to the problem is called decoding. In this paper, each individual has one chromosome, which is encoded by a string of numbers representing the selection priority of a specific project to be completed. Let $E = (e_1, e_2, ..., e_J)$ be a chromosome represented by a string of genes, where J is the number of possible projects to be selected. $e_i, i = 1, 2, ..., J,$ is the ith gene on chromosome E, and its value indicates the selection priority of the ith project. The selection priority for each project is randomly generated within $[1, J]$ exclusively. Thus, to initialize a chromosome (i.e., an individual) is to generate J random numbers within $[1, J]$. An example of a chromosome is shown in Fig. 2.

The main idea of decoding is to choose the one with the highest selection priority value from the candidate set as the successive project to be implemented. In this paper, a connectivity information matrix *Mark[i][n]* is constructed to store whether node *n* is at the end of segment *i* (i.e., project *i*), where $i = 1, 2, ..., J$, and $n = 1, 2, ..., N$. *N* is the number of nodes in the transit network. Besides, a vector *I* is used to indicate whether a node is completed. A procedure to generate a feasible solution to the problem from a chromosome is displayed as follows.

Step 1. Initialize the candidate set by including all the feasible projects.

Step 2. Choose the project with the highest selection priority value from candidate set.

Step 3. Update vector *I* by checking constraints (18)-(19).

Step 4. Update the candidate set by deleting the projects that have been already completed and making changes by checking *Mark* and Constraint (17).

Step 5. Check whether the candidate set is empty. If so, stop and output the sequence of projects to be completed. If not, repeat steps 2-4.

It should be noted that since the values of selection priority for projects are distinct, each chromosome can uniquely determine a feasible sequence of projects. As discussed in the last paragraph in Section 3, a feasible sequence of projects can eventually determine a development plan. Therefore, each chromosome can be uniquely decoded into a feasible solution to the problem. With this genetic encoding scheme, all feasible solutions can be represented by changing the sequence of project priorities.

To further illustrate the process of decoding, we consider a transit network in Fig. 3 and decode the chromosome in Fig. 2 into a feasible solution to this network development problem. At the beginning, initialize the candidate set as (1, 2, 3, 4, 5, 6). Then, choose project 1 from the candidate set as the first project to be implemented due to its highest selection priority, so that the nodes (1, 3) are completed. According to Constraint (17), only projects that connect to segments that have been already completed can be included in the candidate set. Thus, we update the candidate set as (2, 3, 4). Choose project 4 as the successive project because we have 4 (the selection priority of project 4)>3(the selection priority of project 2)>2(the selection priority of project 3). Repeat those steps until the candidate set becomes empty, so that we can obtain a unique feasible sequence of projects as (1, 4, 6, 2, 3, 5).

4.2. Calculating the fitness value

Before calculating the fitness value of an individual, we have to translate a chromosome (e.g., *E*= (6, 3, 2, 4, 1, 5) in Fig. 2) into a feasible sequence of projects (e.g., (1, 4, 6, 2, 3, 5)). In this paper, the fitness function is equal to the value of the objective function as shown in Eq. (16). Therefore, the fitness value of an individual is the discounted total cost of a project sequence. Let ε be the planning horizon. The steps are displayed as follows.

Step 0. Initialization. Let *t* be the counter of periods and set $t = 0$.

Step 1. Calculate the travel demand for OD pairs in period t $\mathbf{q}^{(t)}$ by Eq. (6). Then, determine the daily traffic volume on segments $\mathbf{Q}^{(t)}$ by Eq. (5), headway of transit lines $\mathbf{h}^{(t)}$ by Eq. (15) and Constraint (22), annual user cost $C_u^{(t)}$ by Eq. (7) and annual supplier cost by Eqs. (8)-(12), respectively.

Step 2. Calculate the implementation time of the next project $x^{(t+1)}$ by checking budget constraint in Eq. (23). If $x^{(t+1)} > \varepsilon$, let $x^{(t+1)} = \varepsilon$.

Step 3. Obtain the duration time of period t Δ_t by Eq. (14). Then, calculate the discounted total cost in period t $\Omega^{(t)}$ by Eq. (13) and the cumulative discounted total cost Θ by Eq. (16).

Step 4. If $x^{(t+1)} < \varepsilon$ holds, set $t=t+1$ and go to step 1. Otherwise, stop.

4.3. Selection

Parents are chosen from the population according to a probability which correlates inversely with the fitness value of individuals. To avoid prematurity of the GA, a ranking method proposed by Michalewicz (1996) is adopted. In this method, we first order the individuals in the population from best to worst according to their fitness values, i.e., the individual with the lowest fitness value is the best and is ranked first. Then, we calculate the selection probability of each individual based the exponential ranking value by assuming the lowest fitness value is one. Let p_0 be the selective pressure, which is a positive value between 0 and 1, i.e., $p_0 \in (0,1)$, and p_i be the selection probability of the individual ranked at i. Then, p_i can be expressed as

$$p_i = p_0(1-p_0)^{i-1} \Big/ \left[1-(1-p_0)^M \right], \tag{26}$$

where M is the population size. Next, a roulette wheel approach is used to choose appropriate parents based on their selection probabilities. This process is conducted by spinning the roulette wheel once for each individual in the population. Each time a random number $b \in (0,1)$ is generated, the $i\_$th individual will be selected if $o_{i-1} < b \le o_i$, where o_i is the cumulative probability for each individual.

4.4. Operators

It should be noted that common methods of mutation and crossover are fairly inefficient for our problem since they construct many infeasible solutions with repetitive numbers within one chromosome. To avoid producing such solutions and improve the efficiency, we adopt Partial Matched Crossover (PMX) as the crossover operator and Reciprocal Exchange Mutation (REM) as the mutation operator. These operators are explained by Wang (2001), and thus omitted here.

In general, GA has a strong local search ability, but may get trapped in local optima, which is also known as prematurity. Therefore, the catastrophe mechanism is introduced (Gu et al., 2009). The main idea of this mechanism is to discard the current optima so that the population may produce better

solution. The specific approach in this paper is to regenerate the initial population randomly when the optima stay unchanged over a specified number of generations.

5. Numerical study

In this section, numerical examples are used to illustrate the applications of the proposed model and the contributions of this paper. We consider the urban rail transit network represented in Fig. 3 composed of 3 transit rail lines, 7 nodes (represented by circles) and 6 segments between them. To complete the development of this network, 6 candidate projects are considered. Specifically, each project includes the development of one segment and the two end nodes of this segment (if they are not yet completed). The input data for segments such as length, investment costs and associated rail line are displayed in Table 2. Table 3 shows the daily travel demand between OD pairs. In the following analyses, unless specifically stated otherwise, the input parameters and their baseline values used in the model are the same as those shown in Table 4. We set the planning horizon as 10 years, the annual capital budget as \$250 million and the genetic parameters as follows: population size, *pop_size* = 10; maximum generation, *max_gen* = 100; crossover probability, $P_c = 0.8$; mutation probability, $P_m = 0.5$; the number of implementing catastrophe mechanism, $n_c = 1$. The proposed solution algorithm is coded in MATLAB and run on a ThinkPad Carbon X1 computer with an Intel(R) Core(TM) i5 CPU (2.4 GHz) and 8 GB of RAM. This numerical experiment takes about 0.8 seconds of CPU time.

5.1 Example 1

5.1.1 Optimized solution for rail transit development plan

Table 5 displays the optimized development plan of rail projects and the system performance. It can be seen in Table 5 that 3 projects are selected over a planning horizon of 10 years, i.e., projects 4, 6, and 3, and they are completed sequentially at years 6.00, 8.61 and 9.47, respectively. Over time, the headways of rail lines decrease, but the demand for rail service and discounted cumulative total cost saving increase. Specifically, the headway of Line 1 decreases by 0.13 min from 1.54 min in period 1 to 1.41 min in period 3, and the headway of Line 3 decreases by 0.05 min from 2.91 min in period 2 to 2.86 min in period 3. However, the daily demand for rail service increases by 624.2 thousand from 585.60 thousand riders in period 1 to 1209.80 thousand riders in period 3, and the discounted cumulative total cost saving increases by \$8.04 billion from \$7.21 billion to \$15.25 billion. This occurs because the development of the rail transit network increases the connectivity of OD pairs and hence the demand for rail service, thereby decreasing headways (see Eq. (15)). Thus, the user costs and total costs are reduced and the total cost saving increases.

Fig. 4 shows the changes of the state of the rail transit network over time with the development plan. The bold segments represent those which are already in service in a period. It should be noted that the initial state (from year 0 to 6.00) in which no segments are completed is displayed in Fig. 3. Fig. 4a shows the state of the network in the first period, i.e., from year 6.00 to 8.61. In this period, segment 4 is completed and in service. In period 2 from year 8.61 to 9.47, segment 6 is implemented and connected to segment 4. Both segments 4 and 6 provide rail services, as shown in Fig. 4b. Fig. 4c

54

indicates that segment 3 is completed at the beginning of the third period and in service from year 9.47 to 10. It can be seen from Fig. 4 that throughout the planning horizon, the rail transit network progressively expands to 3 rail lines with a total length of 41 km (i.e., sum of the length of segments 4, 6 and 3).

Fig. 5 shows the changes of discounted cumulative total cost with and without the rail transit investment. It can be seen in Fig. 5 that the total cost curve with investment is under that without investment after year 6.00. This means that the rail transit investment efficiently decreases the total cost of system. It should be noted that in year 6.00, the discounted cumulative total cost with investment is slightly above that without investment due to the capital investment cost of segment 4. Fig. 5 also shows that over the planning horizon, the network development decreases the total cost from $117.53 billion to $102.28 billion.

In order to verify the solution obtained by the proposed GA, we conduct a complete enumeration for the urban transit network shown in Fig. 3. The comparsions of the results are displayed in Table 6. Clearly, the solution obtained by the GA in this paper is consistent with that obtained by complete enumeration. In addition, to test the convergence and stability of the proposed GA, the program is run by 10 times. The results show that each run of the program converges to the same solution. This demonstrates that the proposed GA has good stability. Therefore, we can conclude that the proposed GA is capable of finding a very good and stable solution at acceptable computation cost (i.e., 0.8 seconds vs. 15 seconds).

5.1.2 Sensitivity analysis

To explore the effects of the initial travel demand on the optimized development plan and system performance, we conduct numerical experiments by scaling the basic value of $q_w^{(0)}$ in Eq. (5) by 0.5 down and 1.5 up. Table 7 shows that as the travel demand increases, the number of implemented projects and the total cost saving increases. Specifically, as the initial travel demand increases from $0.5 \times q_w^{(0)}$ to $1.5 \times q_w^{(0)}$, the number of projects selected increases from 2 to 4 and the total cost saving increases from $6.22 billion to $28.62 billion. This is because higher fare revenue can be collected from the operation of completed projects with higher demand, which increases the available budget for network development. Thus, both the number of implemented projects and the total cost saving increase.

Table 8 shows the changes of the optimized development plan with the annual budget level B_0 in Eq. (24). It can be noted in Table 8 that the annual budget level has a significant effect on the optimized development plan and system performance in terms of the number of projects selected, the time of implementation and the total cost saving. Specifically, as the annual budget level increases from $0.8 \times B_0$ to $1.2 \times B_0$, the number of projects selected increases by 3 from 1 to 4, the first investment time decreases by 4 years from year 9.00 to year 5.00 and the total cost saving increases by $22 billion from $0.95 billion to $22.95 billion. This implies that a higher budget level can accelerate the development process and save more costs.

55

5.2 Example 2

To further illustrate the applications of the proposed model and test the performance of the GA on a more complex problem, we apply the proposed model to the rail transit network development of Wuhan city in China. As shown in Fig. 6a, there are 3 rail lines represented by three colors: blue for Line 1, purple for Line 2 and green for Line 4. A rail transit network with 14 nodes (represented by circles) and 13 segments between them is considered, as shown in Fig. 6b. Similarly, we consider the development of one segment and its two end nodes (if they are not yet completed) as a candidate project. The input data for segments and OD pairs are displayed in Tables 9 and 10, respectively. The base values of the input parameter are shown in Table 4. We set the planning horizon as 15 years and the annual budget flow as \$1 billion. The genetic parameters are: population size, *pop_size* = 50; maximum generation, *max_gen* = 500; crossover probability, $P_c = 0.8$; mutation probability, $P_m = 0.2$; the number of implementing catastrophe mechanism, $n_c = 20$; and run 10 times. This numerical experiment requires an average CPU time of about 13 min. Using the proposed GA, we can obtain the same solution for all runs, which shows that the proposed GA maintains its stability on a more complex problem.

The optimized development plan and headways of rail lines are displayed in Table 11. It can be seen that 11 projects are developed over a planning horizon of 15 years with a total cost of \$99.28 billion. Specifically, projects 3, 6, 8, 10, 4, 2, 5, 9, 11, 7 and 1 are completed in sequence at years 0.29, 1.60, 3.63, 4.67, 6.45, 7.98, 10.17, 11.40, 12.06, 12.76, 13.86, respectively. This result is roughly consistent with the realistic development of the urban rail transit network in Wuhan between 2000 and 2014, as shown in Fig. 1.

Since the enumeration of this problem with 13 candidate projects (i.e. 13! possible solutions) requires extensive computation time, and no existing method can guarantee a globally optimal solution, it is difficult to verify the solution obtained by the proposed GA. In this paper, a statistical method is adopted to evaluate the solution (as in Jong and Schonfeld, 2003 or Shayanfar et al., 2016). The main steps are as follows. First, a large sample of solutions is randomly generated. These solutions should be representative and independent of each other to ensure the generality of the sample. Then, the fitness values of the solutions in the sample are calculated. Next, a distribution is fitted to the fitness values and checked with Chi-Square or K-S tests. It should be noted that the fitted distribution should approximate the actual distribution of fitness values for all possible solutions in the search space due to the representativeness and randomness of the sample. Finally, the cumulative probability of the solution in the distribution can be calculated. This cumulative probability represents the probability that is the other solutions in the distribution smaller than the obtained solution. Therefore, the lower the probability, the better the solution.

In this paper, a sample size of 100,000 independent solutions is randomly generated, for which the minimum of the fitness values is 99.88×10^9 and the maximum is 131.21×10^9. Note that the best solution found by the proposed GA is 99.28×10^9 which is better than any of the 100,000 randomly generated solutions, as shown in Fig. 7. The distribution of the fitness values for the solutions in the sample is supposed to cover the fitness values for all possible solutions in the search space. Actually, it does not. This means that better solutions (i.e., having lower fitness values) are extremely rare for this example and are unlikely to be included in a randomly generated sample. The best fitting distribution among those searched is the generalized extreme value distribution, i.e., GEV($\mu = 112.514 \times 10^9$,

56

$\sigma = 5.27436$, $\kappa = -0.145511$), as is shown in Fig. 7. Its probability density function can be expressed as

$$f(x) = \frac{1}{\sigma} \varphi(x)^{\kappa+1} e^{-\varphi(x)}, \text{ where } \varphi(x) = \begin{cases} \left(\left(1 + \kappa \left(\frac{x-\mu}{\sigma} \right) \right)^{-1/\kappa}, \text{ if } \kappa \neq 0, \\ e^{-(x-\mu)/\sigma}, \text{ if } \kappa = 0. \end{cases} \tag{27}$$

The cumulative probability of the best solution found by the proposed GA (i.e., 99.28×10^9 in Table 11) can be calculated by integrating $f(x)$ from 0 to 99.28×10^9. The result is 2.0552×10^{-4}, which means that the solution obtained by the proposed GA dominates 99.98% of the solutions in the distribution, as well as 100% of the 100,000 randomly generated solutions. That is to say, the best solution found, although not guaranteed to be globally optimal, is still remarkably good when compared with other possible solutions in the search space. This suggests that the accuracy of the proposed development scheduling method is limited far more by the accuracy of input data than by the optimization capability of the GA.

6. Conclusions and further studies

To address the dynamic development problem of urban rail transit networks with limited budgets, this paper proposes a novel model to optimize the development plan of rail transit projects over a planning horizon. The proposed model determines which projects should be implemented and when to complete these projects together with train headways by minimizing the present value of the total cost. The time-varying demand and the interrelation among projects are explicitly considered. Specifically, the model captures how the travel demand for rail service, the headway of rail lines and the network development decision change over time. In this dynamic decision making process, the budget constraint is reformulated to include possible internal funding, such as the fare revenue generated from the operation of the transit rail lines. The reformulated budget constraint reflects interrelations among projects in the transit network since the capital used for development is partly supplied by fare revenue collected from the rail operation. A GA approach is designed to solve the problem, and the properties of the solution found by the proposed GA are verified.

Results show that (i) the GA approach developed here is capable of finding a quite good and stable solution at acceptable computaion cost. (ii) The development of the rail transit network can significantly increase the demand for rail service and reduce the total cost. (iii) Higher travel demand can encourage more intensive network development and increase the total cost saving. This helps explain why many large cities in China such as Beijing and Shanghai are investing heavily in transit development. (iv) A higher budget level can accelerate the development process over the planning horizon and reduce total costs. The proposed model can serve as a useful tool for making development plan of transit networks from an economic viability and cost-effectiveness perspective.

Although this paper provides a new venue for addressing the transit network development problem, some further extensions seem worth pursuing:

1. Travel demand is assumed to be attracted to rail service when OD pairs are connected by rail lines, but is not affected by the transit service characteristics. However, travelers are usually sensitive to the travel cost and thus the transit service level (Li et al., 2012a; Peng et al., 2017). Therefore, it

seems desirable to extend the proposed model to capture the responses of passengers to the quality of the rail transit line service.

2. In this paper, the proposed model is deterministic because the demand and supply sides are assumed to be deterministic. However, in reality there are various random factors (e.g., inflation and economic changes) which affect the investment of rail lines and the operations of rail services. It is thus especially important for the authority to consider the investment and operational risks of rail transit projects in the development issue of urban rail transit networks, which is left for our future study.

3. This paper focuses mainly on rail mode, and neglects the competition and substitution effects between private auto and transit modes. It seems desirable to extend the proposed models to consider different modes and analyze the transit network development in a multi-modal transportation system (Li et al., 2012b; Ma and Lo, 2013).

4. Urban spatial structure in terms of households' residential location choices and housing market has a direct effect on travel demand pattern (Li et al., 2012c; Li and Peng, 2016; Wang and Lo, 2016; Ng and Lo, 2017), and thus on the rail transit service and the network development process. Therefore, it seems worthwhile to extend the proposed model to explore the effects of urban spatial structure on transit network development.

Acknowledgments

References

Cheng, W.C., Schonfeld, P, 2015. A method for optimizing the phased development of rail transit lines. Urban Rail Transit, 1(4), 227-237.

CURTA (Chinese Urban Rail Transit Association), 2017. Annual Urban Rail Transit Statistics and Analysis Report in 2016. <http://www.camet.org.cn/index.php?m=content&c=index &a=show&catid=18&id=1047>.

Farahani, R.Z., Miandoabchi, E., Szeto, W.Y., Rashidi, H., 2013. A review of urban transportation network design problems. European Journal of Operational Research, 229(2), 281-302.

Jong, J.C., Schonfeld, P., 2003. An evolutionary model for simultaneously optimizing three-dimensional highway alignments. Transportation Research Part B, 37(2), 107-128.

Gao, Z., Sun, H., Shan, L.L., 2004. A continuous equilibrium network design model and algorithm for transit systems. Transportation Research Part B, 38(3), 235-250.

Gu, J., Gu, X., Gu, M., 2009. A novel parallel quantum genetic algorithm for stochastic job shop scheduling. Journal of Mathematical Analysis and Applications, 355(1), 63-81.

Li, C., Yang, H., Zhu, D., Meng, Q., 2012. A global optimization method for continuous network design problems. Transportation Research Part B, 46(9), 1144-1158.

Li, Z.C., Lam, W.H., Wong, S.C., Sumalee, A., 2012a. Design of a rail transit line for profit maximization in a linear transportation corridor. Transportation Research Part E, 48(1), 50-70.

Li, Z.C., Lam, W.H.K., Wong, S.C., 2012b. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. Transportation Research Part B, 46 (1), 30- 49.

Li, Z.C., Lam, W.H., Wong, S.C., Choi, K., 2012c. Modeling the effects of integrated rail and property development on the design of rail line services in a linear monocentric city. Transportation Research Part B, 46(6), 710-728.

Li, Z.C., Peng, Y.T., 2016. Modeling the effects of vehicle emission taxes on residential location choices of different-income households. Transportation Research Part D, 48, 248-266.

Liu, H., Wang, D.Z., 2015. Global optimization method for network design problem with stochastic user equilibrium. Transportation Research Part B, 72, 20-39.

Luathep, P., Sumalee, A., Lam, W.H., Li, Z.C., Lo, H.K., 2011. Global optimization method for mixed transportation network design problem: a mixed-integer linear programming approach. Transportation Research Part B, 45(5), 808-827.

Ma, X., Lo, H.K., 2013. On joint railway and housing development strategy. Transportation Research Part B, 57, 451- 467.

Michalewicz, Z., 1996. Evolution strategies and other methods. In Genetic algorithms+ data structures= evolution programs (pp. 159-177). Springer, Berlin, Heidelberg.

Ng, K.F., Lo, H.K., 2017. On joint railway and housing development: Housing-led versus railway-led schemes. Transportation Research Part B, 106, 464-488.

Peng, Y.T., Li, Z.C., Choi, K., 2017. Transit-oriented development in an urban rail transportation corridor. Transportation Research Part B, 103, 269-290.

Shayanfar, E., Abianeh, A.S., Schonfeld, P., Zhang, L., 2016. Prioritizing interrelated road projects using metaheuristics. Journal of Infrastructure Systems, 22(2), 04016004.

Sun, Y., Guo, Q., Schonfeld, P., Li, Z., 2017. Evolution of public transit modes in a commuter corridor. Transportation Research Part C, 75, 84-102.

Sun, Y., Schonfeld, P., Guo, Q., 2018. Optimal extension of rail transit lines. International Journal of Sustainable Transportation. https://doi.org/10.1080/15568318.2018.1436730.

Wang, D.Z., Liu, H., Szeto, W.Y., 2015. A novel discrete network design problem formulation and its global optimization solution algorithm. Transportation Research Part E, 79, 213-230.

Wang, D.Z., Lo, H.K., 2016. Financial sustainability of rail transit service: The effect of urban development pattern. Transport Policy, 48, 23-33.

Wang, S., Meng, Q., Yang, H., 2013. Global optimization methods for the discrete network design problem. Transportation Research Part B, 50, 42-60.

Wang, S.L., 2001. Simulation and optimization of interdependent waterway improvement projects (Doctoral dissertation, University of Maryland, College Park).

Wang, S.L., Schonfeld, P., 2008. Scheduling of waterway projects with complex interrelations. Transportation Research Record, 2062, 59-65.

Yin, Y., Li, Z.C., Lam, W.H., Choi, K., 2014. Sustainable toll pricing and capacity investment in a congested road network: a goal programming approach. Journal of Transportation Engineering, 140(12), 04014062.

Zhang, L., Yang, H., Wu, D., Wang, D., 2014. Solving a discrete multimodal transportation network design problem. Transportation Research Part C, 49, 73-86.

Table 1 Contributions to transportation infrastructure investment models.

Citation	Type of infrastructure	Considering time horizon or not	Considering interrelation among projects or not
Wang et al. (2013)	Road network	×	×
Li et al. (2012)	Road network	×	×
Luathep et al. (2011)	Road network	×	×
Gao et al. (2004)	Transit network	×	×
Sun et al. (2018)	Rail line	√	×
Shayanfar et al. (2016)	Road network	√	√
This paper	Transit network	√	√

Note: "√" means that the associated item is considered, whereas "×" means that the associated item is not considered.

Table 2 Input data for segments.

Segment No.	Segment length (km)	Segment investment costs (million $)	Associated rail line
1	12	1250	2
2	10	1050	1
3	8	850	2
4	15	1500	1
5	9	950	1
6	18	1800	3

Table 3 Daily travel demands between OD pairs (thousands person trips).

Nodes No. (O/D)	1	2	3	4	5	6	7
1	0	10	30	11	12	24	25
2	10	0	35	10	27	20	12
3	30	35	0	30	40	25	20
4	11	10	30	0	30	10	15
5	12	27	40	30	0	35	20
6	24	20	25	10	35	0	15
7	25	12	20	15	20	15	0

Table 4 Input parameters for numerical examples.

Symbol	Definition	Baseline value
λ_1	Value of in-vehicle time ($/h)	15
λ_2	Value of waiting time ($/h)	30
V	Average speed of trains (km/h)	40
f	Marginal fare by transit ($/km)	0.2
g_1	Base growth rate of travel demand (year)	0.02

g_2	Annual growth rate caused by network development (year)				0.03
ρ	Average number of days of travel per traveler per year				250
η	Marginal maintenance cost of transit lines (million \$/km/year)				5
β	Annual operating cost per train (million \$/year)				3
r	Discount rate				0.05
τ	Peak-hour factor				0.1
K_{veh}	Capacity of vehicles (passengers/vehicle)				1500
p_0	Selective pressure				0.2

Table 5 Optimized development plan for rail transit network and resulting system performance.

Period No.	Segment developed	Completion time (year)	Train headways of line 1, 2 and 3 (min)			Daily demand for rail service (thousand person trips)	Discounted cumulative total cost saving (billion \$/year)
			h_1	h_2	h_3		
1	4	6.00	1.54	-	-	585.60	7.21
2	6	8.61	1.44	-	2.91	930.27	11.71
3	3	9.47	1.41	2.12	2.86	1209.80	15.25

Notes: (1) The completion time of projects is also the starting or ending time of periods. (2) The discounted cumulative total cost saving is calculated by the discounted cumulative total cost without investment minus the that with investment.

Table 6 Comparisons of results obtained by GA and complete enumeration.

GA (computation time: 0.8 seconds)			Complete enumeration (computation time: 15 seconds)		
Period No.	Segment developed	Completion time (year)	Period No.	Segment developed	Completion time (year)
1	4	6.00	1	4	6.00
2	6	8.61	2	6	8.61
3	3	9.47	3	3	9.47

Table 7 Effects of travel demand on the optimized development plan and system performance.

	0.5×base value	Base value	1.5×base value
Number of developed projects	2	3	4
Developed projects	4 (6.00)	4 (6.00)	4 (6.00)
(completion time, year)	3 (7.81)	6 (8.61)	6 (7.98)
		3 (9.47)	2 (8.76)
			3 (9.30)
Total cost saving (billion \$)	6.22	15.25	28.62

Table 8 Effects of annual budget on the optimized development plan and system performance.

	0.8×base value	Base value	1.2×base value
Number of developed projects	1	3	4
Developed projects (completion time)	6 (9.00)	4 (6.00)	4 (5.00)
		6 (8.61)	6 (7.46)
		3 (9.47)	2 (8.49)
			3 (9.20)
Total cost saving (billion $)	0.95	15.25	22.95

Table 9 Input data for segments of Wuhan rail transit network.

Segment No.	Segment length (km)	Segment investment costs (million $)	Associated rail line
1	8.4	2280	1
2	9.8	2325	1
3	3.9	292.5	1
4	10.3	2475	2
5	20.0	3600	2
6	9.2	1380	1
7	6.6	1387.5	1
8	8.0	2400	2
9	12.5	2250	4
10	9.1	1350	2
11	4.6	1275	4
12	10.9	2500	4
13	5.5	990	4

(*Sources: http://www.whrt.gov.cn/ and Baidu Map*)

Table 10 Initial daily travel demands between OD pairs of Wuhan rail transit network

(thousand person trips).

| Nodes No. (O/D)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	3.2	2.4	6	2.4	3.2	6	3.2	2	1.6	1.6	0.8	0.4	0.4
2	3.2	0	8	10	3.2	4	9.6	3.2	3.2	2.4	2.4	1.6	0.8	0.4
3	2.4	8	0	16	8	8	20	10	4	3.2	3.2	2.4	1.6	0.8
4	6	10	16	0	6.4	10	16	12	8	4.8	4.8	2.4	1.6	1.2
5	2.4	3.2	8	6.4	0	9.6	4.8	4	6.4	4.8	4.8	2.4	1.6	0.8
6	3.2	4	8	10	9.6	0	2.4	2.4	3.2	3.2	3.2	1.6	1.2	0.8
7	6	9.6	20	16	4.8	2.4	0	1.6	3.2	2.4	16	1.6	1.2	0.8
8	3.2	3.2	10	12	4	2.4	1.6	0	2.4	1.6	1.6	1.2	0.8	0.8
9	2	3.2	4	8	6.4	3.2	3.2	2.4	0	4	32	2.4	2.4	1.6
10	1.6	2.4	3.2	4.8	4.8	3.2	2.4	1.6	4	0	8.8	4.8	3.2	2.4
11	1.6	2.4	3.2	4.8	4.8	3.2	16	1.6	32	8.8	0	6.4	1.6	1.2
12	0.8	1.6	2.4	2.4	2.4	1.6	1.6	1.2	2.4	4.8	6.4	0	0.8	0.4

| 13 | 0.4 | 0.8 | 1.6 | 1.6 | 1.6 | 1.2 | 1.2 | 0.8 | 2.4 | 3.2 | 1.6 | 0.8 | 0 | 2.4 |
| 14 | 0.4 | 0.4 | 0.8 | 1.2 | 0.8 | 0.8 | 0.8 | 0.8 | 1.6 | 2.4 | 1.2 | 0.4 | 2.4 | 0 |

Table 11 Optimized network development plan and headways of rail lines in Wuhan.

Period No.	Segment developed	Completion time (year)	Train headways of line 1, 2 and 4 (min)			Daily demand for rail service (thousand person trips)	Discounted cumulative total cost (billion \$/year)
			h_1	h_2	h_3		
1	3	0.29	1.38	-	-	292.89	19.04
2	6	1.60	1.59	-	-	564.44	38.78
3	8	3.63	1.51	2.09	-	862.10	47.77
4	10	4.67	1.46	1.94	-	1079.08	60.89
5	4	6.45	1.38	1.28	-	1353.40	70.51
6	2	7.98	1.04	1.22	-	1608.85	82.04
7	5	10.17	0.96	0.97	-	1869.98	87.57
8	9	11.40	0.91	0.92	3.90	2080.53	90.22
9	11	12.06	0.89	0.90	3.07	2277.75	92.85
10	7	12.76	0.78	0.87	3.03	2460.48	96.71
11	1	13.86	0.69	0.83	2.96	2672.02	99.28
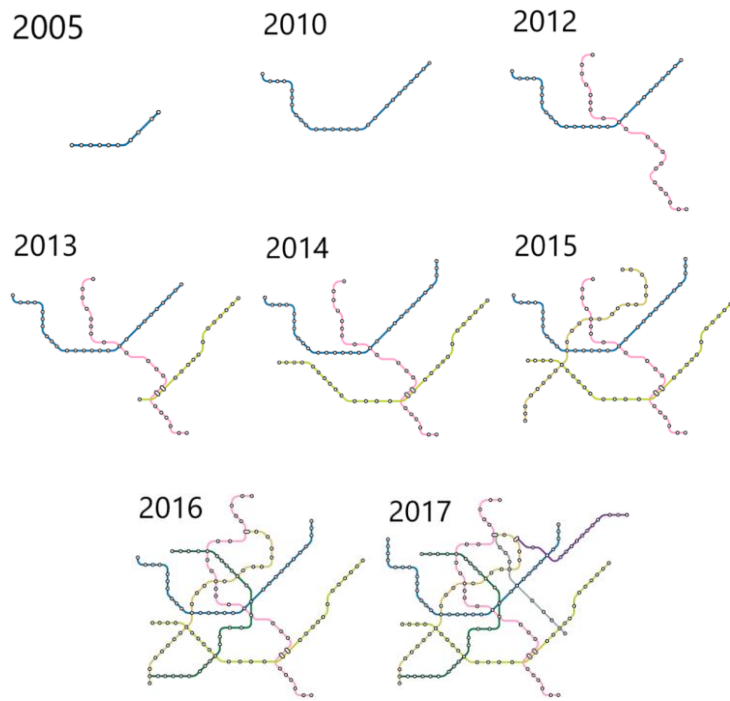
Fig.1. Development process of rail transit network in Wuhan, China

(*Sources: http://www.whrt.gov.cn/ and https://en.wikipedia.org/wiki/Wuhan_Metro*).

Project ID	1	2	3	4	5	6
Selection priority of projects	6	3	2	4	1	5

Fig. 2. Example of a chromosome.
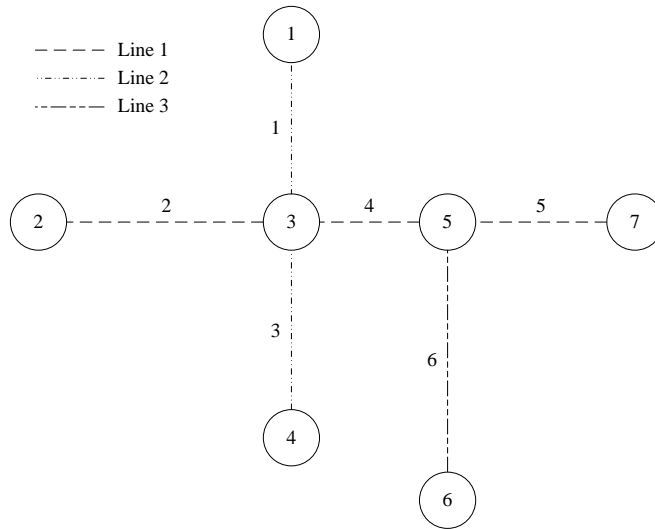
Fig. 3. Example of an urban rail transit network.

(a) t=1(year 6.00- 8.61)   (b) t=2 (year 8.61- 9.47)   (c) t=3 (year 9.47- 10)
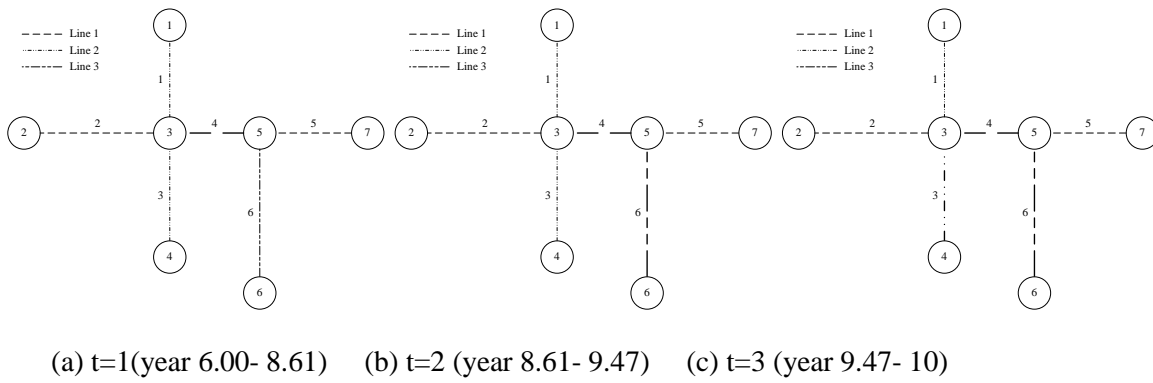
Fig. 4. Evolution of the state of the rail transit network
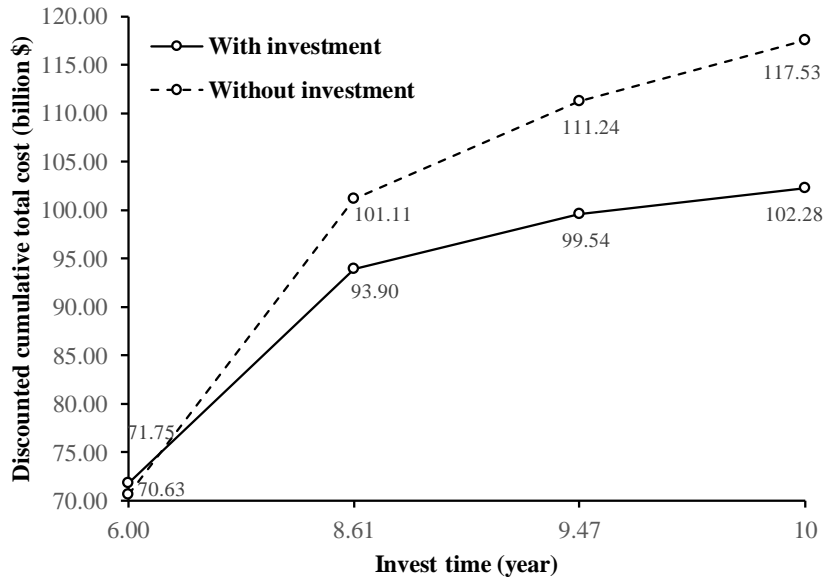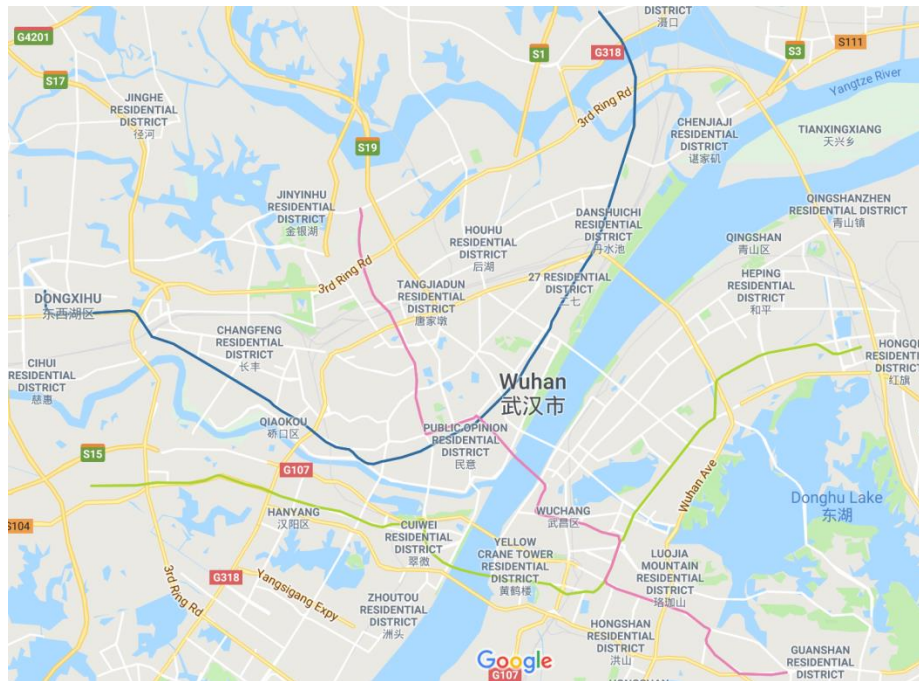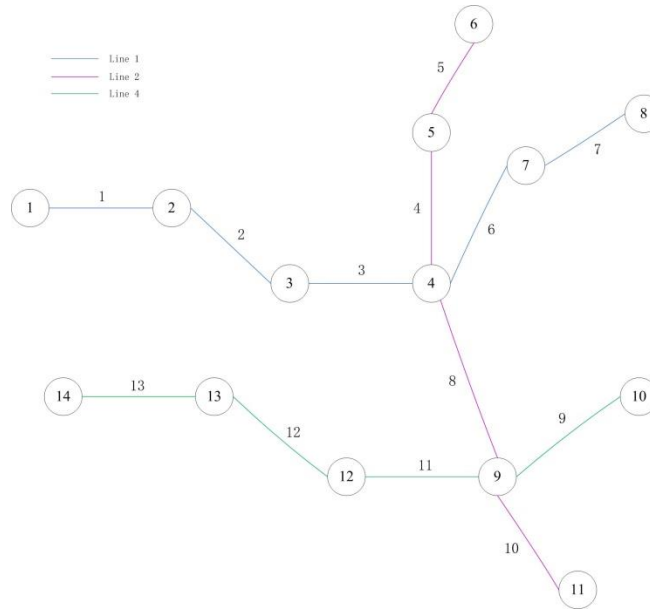
Fig. 5. Changes of discounted cumulative total cost with and without the rail investment.
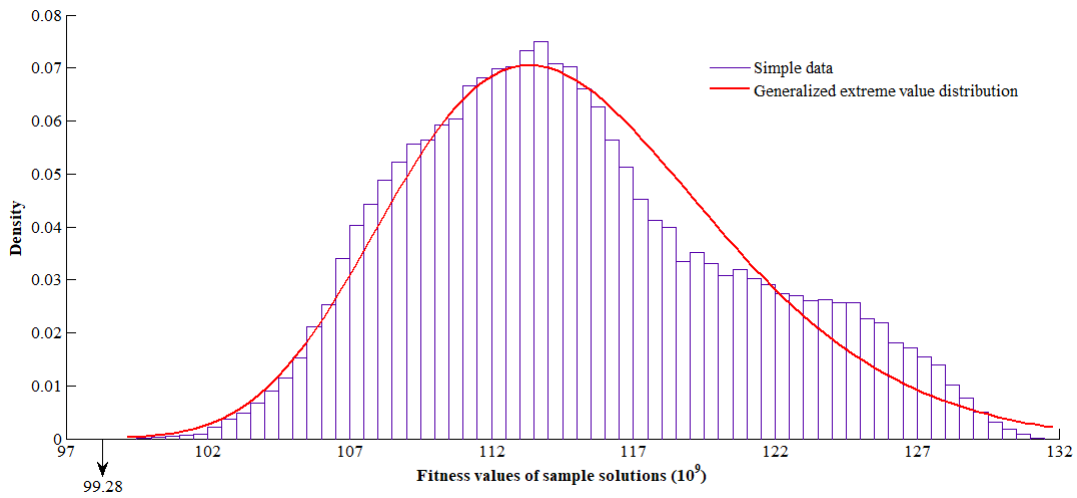

(a)

(b)

Fig. 6. Map of Wuhan subway lines (blue for Line 1, purple for Line 2 and green for Line 4): (a) urban rail transit network of Wuhan, China; (b) candidate rail transit projects.

Fig. 7. Fitted generalized extreme value distribution of the fitness values of the sample.