

A Method for Optimizing the Phased Development of Rail Transit Lines

Wei-Chen Cheng¹ · Paul Schonfeld²

Received: 11 November 2015 / Revised: 23 December 2015 / Accepted: 29 December 2015 / Published online: 21 January 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This paper develops a method for optimizing the construction phases for rail transit line extension projects with the objective of maximizing the net present worth and examines the economic feasibility of such extension projects under various financial constraints (i.e., unconstrained, revenue-constrained, and budget-constrained cases). A Simulated Annealing algorithm is used for solving this problem. Rail transit projects may be divided into several phases due to budget limits or demand growth that justifies different sections at different times. A mathematical model is developed to optimize these phases for a simple, one-route rail transit system, running from a Central Business District (CBD) to a suburban area. Some interesting results indicate that the economic feasibility of links with low demand is affected by the completion time of those links and their demand growth rate after their implementation. Sensitivity analysis explores the effects of interest rates on optimized results (i.e., construction phases and objective value). With further development, such a method should be useful to transportation planners and decision-makers in optimizing construction phases for rail transit line extension projects.

Keywords Rail transit · Phased development · Optimization · Simulated annealing · Net present worth

1 Introduction

Projects for new or extended rail transit lines may be subdivided into phases based on demand growth considerations and budget limits over time. Any additions of stations or extensions of rail lines affect many users and involve substantial investments. Consequences of adding stations may include increased mobility, higher land values, increased employment opportunities, environmental impacts, and reduced congestion. Therefore, such a project requires a comprehensive evaluation of all direct and indirect consequences, including positive and negative effects on different affected groups [22]. No general guidelines are yet available on how many phases are needed and when each phase should be implemented. The phases and execution time are usually based on available budgets, demand forecasts, and political reasons (e.g., equity among regions). The scheduled phases may be far from optimal if significant effects of extensions are neglected, such as faster demand growth after service quality and accessibility improvements (e.g., new stations). Scheduling decisions affect system performance over the entire analysis period. Therefore, in order to overcome existing analytic weaknesses, we propose a model that optimally subdivides a predetermined rail transit line into sections for phased development and optimizes the implementation times of those sections over a planning horizon. The evaluation and scheduling of additions to lines (i.e., links and stations) are performed jointly by this model. Based on various specified evaluation criteria, the model can optimize a phased development plan. This

✉ Paul Schonfeld
pschon@umd.edu
Wei-Chen Cheng
rcheng@rkk.com

¹ Rummel Klepper & Kahl, L.L.P., 81 Mosher St., Baltimore, MD 21217, USA

² Department of Civil & Environmental Engineering, University of Maryland, College Park, MD 20742, USA

model is demonstrated here for one hypothetical rail transit line but is designed to be applicable to any such transit lines.

Tavares [21] optimizes the schedule for a set of interconnected railway projects with the purpose of maximizing the net present worth (NPW), using Dynamic Programming. This model is applicable for scheduling large sets of expensive and interconnected development projects under tight capital constraints and with a marginal net present value. He notes that maximizing the NPW of a project in terms of its schedule under eventual restrictions concerning its total duration can be considered as a dual perspective of the problem of minimizing makespan (defined as the total duration of a project) with resource constraints. The model presented in the paper does not consider demand reductions during construction. The items considered in NPW are only construction expenditures and payments received after completion of projects. Since it is a renewal project, all the items that are affected by the project should be taken into account.

Kolisich and Padman [9] summarize and classify previous studies on the resource-constrained project scheduling problem (RCPSP) by their objectives and constraints: NPW maximization and makespan minimization, with and without resource constraints. For the resource-unconstrained case, generally it is optimal to schedule jobs with associated positive cash flows as early as possible, and jobs with net negative cash flows as late as possible, subject to restrictions imposed by network structure.

Matisziw et al. [14] propose an optimization model to determine route extension networks for bus transit systems. It is similar to a routing problem that maximizes covering areas and minimizes the extension length under resource constraints. It is important to expand the existing service network to tap into emerging areas of demand not being served. Maximizing network coverage can increase ridership. While increasing this potential ridership is significant, it is necessary to keep any route extension to a minimal length. Extending routes to low-demand areas could result in low service utilization. In our present study, the NPW maximization objective determines how far routes should be extended to low-density suburbs.

Wang and Schonfeld [23] develop a simulation model to evaluate waterway system performance and optimize the improvement project decisions with demand model incorporated. They argue that minimizing total costs rather than maximizing the NPW over the entire analysis period is not valid in a system where demand is elastically affected by the system improvements being optimized. The results show how demand elasticity can be used in estimating net benefits. Shayanfar et al. [16] compare the relative merits of three metaheuristic algorithms, namely simulated annealing, tabu search, and a

genetic algorithm, for selecting and scheduling improvements in road networks.

Numerous other researchers have developed related models for optimizing various characteristics of public transportation systems. These include Guan et al., [5], Fan and Machemehl [3], Zhou et al. [24], Li et al. [11], Tsai et al. [20], DiJoseph and Chien [2], Kim and Schonfeld [8], and Markovic et al. [13]. Kim et al. [7] optimized vertical alignments and speed profiles for rail transit lines. Lai and Schonfeld [10] optimized the location of rail transit lines and stations, based on GIS databases and using a genetic algorithm, but without considering phasing decisions. Lo and Szeto [12] and Szeto et al. [19] deal with the timing of improvements in discrete network design. Guihaire and Hao [6] review transit network design and scheduling approaches, while Farahani et al. [4] review urban transportation network design more generally.

The modeling approach used in our present study is partly based on a model of Chien and Schonfeld [1], except for the decision variables. They developed a model that jointly optimized the characteristics of a rail transit route and its associated feeder bus routes in order to minimize total costs. Somewhat similarly, Spasovic and Schonfeld [17] also optimize the transit service coverage with a minimum total cost objective. Their analytic results showed that in order to minimize total costs, the operator cost, user access cost, and user wait cost should be equalized. They also noted that the most significant factor in determining the rail line length is the demand. Thus, no route completion constraint is considered in our present model because it might overextend routes into distant suburbs with insufficient demand density. Sun and Schonfeld [18] analyze a related phased development problem, but for airport facilities rather than rail transit lines.

Although the published studies we found do not deal with the optimized phased development of transit lines, this problem can be treated as an RCPSP with unique characteristics. First, the activities in this project represent the stations to be added. Second, the precedence relations in this problem are much easier than those in the general project scheduling problem. Our transit line can only be extended sequentially from one end (i.e., CBD) to the other (We can still treat a line through a CBD as two end-to-end radial lines). Third, constraints on both capital budget and revenue are considered in this study. For the capital budget constraint, subsidies are equally distributed here within each given time interval, although any distribution may easily be specified. The revenue constraint is used for balancing the operational expenditure. It is important to note that the resource constraints vary over the entire time horizon, since these two constraints are affected by the operational situation and decision made in previous years. Hence, this problem is a dynamic RCPSP. NPW maximization is our chosen objective. All the quantifiable items

that would be affected by the extension should be considered in this problem (e.g., user waiting costs, in-vehicle costs, and operating and maintenance costs), including socio-economic effects if they can be quantified and estimated correctly. Due to the complexity of the dynamic RCPSP, including the pervasiveness of local optima, we use a Simulated Annealing algorithm to solve this problem. The model formulation and design of the SA algorithm are presented below.

2 Model Formulation

Table 1 defines the notation used in the paper. The following simplifying assumptions are made here.

A given demand at the starting time interval ($t = 0$) is already consistent with network equilibrium.

1. Transit routes and station locations are predetermined. Hence, user access costs are omitted from this analysis.
2. Effects of development schedules of other transportation system changes on the demand of our line are neglected.
3. Stations can only be added sequentially from the CBD outward. With a double crossover track at every station, any station can be at least temporarily the line's terminal station. Hence, turnaround time is omitted from this analysis.
4. There are no binding construction time constraints.
5. Potential demand for each O/D pair increases at a higher rate after the station is completed.

Table 1 Notation

Variables	Descriptions	Units
B	Total benefit	\$
C	Total cost	\$
C_C	Capital cost	\$
C_I	In-vehicle cost	\$
C_M	Maintenance cost	\$
C_O	Operating cost	\$
C_S	Supplier cost	\$
C_U	User cost	\$
C_W	Waiting cost	\$
d	Station spacing	mile
f	Taxation ratio for covering operational expenditure	%
F_T	Fleet size	vehicle
h	Headway	h
i	The origin in the O/D matrix	–
j	The destination in the O/D matrix	–
k	Capital cost for station and rail line	\$
m	The row in the O/D matrix	–
n_C	Number of cars needed per train	cars/vehicle
P	Demand function	–
NPW	Net present worth of total benefits	\$
q_{ij}	Rail passenger flows from origin i to destination j	people
r	Demand growth rate	–
R	Round trip time	h
s	Interest rate	–
t	Time interval	–
t_d	Dwell time	hour
u_I	Unit cost of user in-vehicle time	\$/passenger-h
u_L	Maintenance unit cost	\$/passenger-mile
u_T	Hourly operating cost	\$/vehicle-h
u_W	Unit cost of user waiting time	\$/passenger-h
U_B	User benefit	\$
V	Cruise speed	miles/h
y	Decision variable	–

6. Capital costs are reduced if multiple stations and their links are built together.
7. The interest rates are effective rates which already consider inflation.

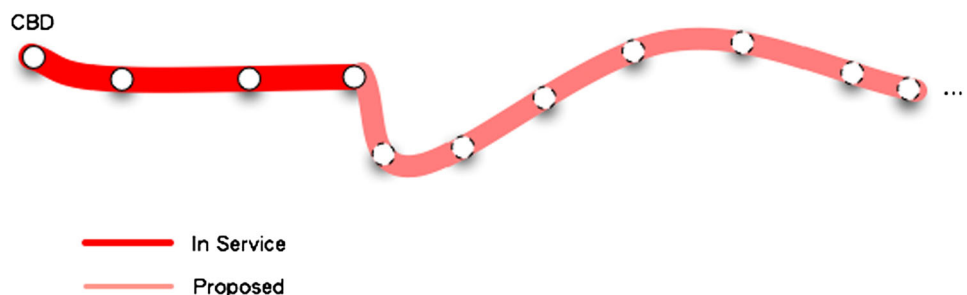
Figure 1 shows the proposed example rail transit line, which is 54.4 miles long and has 30 stations. Currently, only 4 stations are completed and in service. The study’s time horizon is 30 years. Our binary decision variable $y_i^{(t)} = 1$ if link i and its station already exist in time period t ; $y_i^{(t)} = 0$ if link i is yet unbuilt in time period t . Here link i is defined as the section between stations $i - 1$ and i , and link i includes station i . The first time $y_i^{(t)}$ changes from 0 to 1 which indicates that link i is added in year t . In the long term, the traffic increase may occur due to demographic and economic growth. Demand growth is considered here by multiplying the demand relation for the initial period ($t = 0$) with a compound growth rate $(1 + r)^t$, where r is the growth rate per time interval (e.g., per week, month, or year) and t represents intervals of growth (Fig. 2). The baseline demand function for each origin/destination pair is a linear demand function (i.e., $Q = a - b * P$).

As discussed above, the origin/destination (O/D) matrix values can continuously increase at a specific annual growth rate based on traffic demand forecasts. $q_{ij}^{(t)} = -q_{ij}^{(0)} \times (1 + r)^t, \forall i, j$, where q_{ij} denotes rail passenger flows from origin i to destination j . For our numerical study, the O/D matrix is symmetric, with $q_{ij} = q_{ji}$. There are 4 stations in service in time interval zero. The O/D matrix is

$$D^{(t)} = \begin{bmatrix} - & y_2q_{12} & y_3q_{13} & y_4q_{14} & y_5q_{15} & y_6q_{16} & \dots \\ y_2q_{21} & - & y_3q_{23} & y_4q_{24} & y_5q_{25} & y_6q_{26} & \dots \\ y_3q_{31} & y_3q_{32} & - & y_4q_{34} & y_5q_{35} & y_6q_{36} & \dots \\ y_4q_{41} & y_4q_{42} & y_4q_{43} & - & y_5q_{45} & y_6q_{46} & \dots \\ y_5q_{51} & y_5q_{52} & y_5q_{53} & y_5q_{54} & - & y_6q_{56} & \dots \\ y_6q_{61} & \dots & \dots & \dots & \dots & - & \dots \\ y_7q_{71} & \dots & \dots & \dots & \dots & \dots & - \end{bmatrix}^{(t)},$$

where at $t = 0, y_1 = y_2 = y_3 = y_4 = 1, y_5 = y_6 = \dots = 0$.

Fig. 1 Proposed route



2.1 Benefit Function

User benefit (U_B), in any time interval t , is defined as the area under the demand (=marginal user benefit = P) curve for that interval, integrated from 0 to $q_{ij}^{(t)}$, where $q_{ij}^{(t)}$ is the traffic flow from i to j in the t th simulation interval (Fig. 2). Since q_{ij} may fluctuate in different intervals, the overall user benefit for the entire analysis period is

$$U_B = \sum_t \sum_i \sum_j \left(\int_0^{q_{ij}^{(t)}} P \cdot dQ \right), \quad i \neq j. \tag{1}$$

2.2 Cost Function

The user cost (C_U) consists of three components: in-vehicle cost, waiting cost, and access cost. Access cost is the total demand multiplied the access time. Because we assume that station locations are predetermined, the access cost might be omitted. The waiting cost, C_W , is the total demand multiplied by the waiting time (which is approximated as half of the headway, h), and the unit cost of user waiting time, u_W (\$/passenger-h):

$$C_W^{(t)} = D^{(t)} \times \frac{h}{2} \times u_W. \tag{2}$$

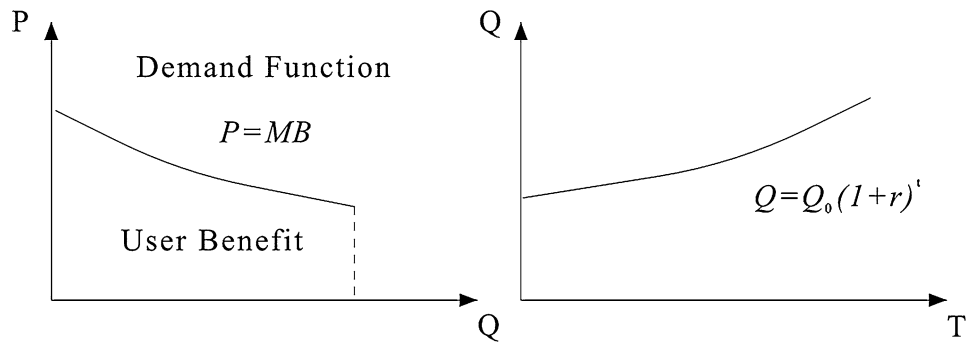
In-vehicle cost, C_I , is the through flow multiplied by the in-vehicle time which includes the riding and dwell time and the unit cost of in-vehicle time, u_I (\$/passenger-hour). Through flow is equal to inflow minus outflow at each station, and it can be determined from the O/D matrix:

$$\text{Through flow} = 2 \times \sum_{m=1}^{30} \sum_{i=1}^m \left(\sum_{j=i+1}^{30} y_jq_{ij} - \sum_{j=1}^i y_iq_{ij} \right), \tag{3}$$

where m is the row in the O/D matrix, i is the origin in the O/D matrix, j is the destination in the O/D matrix.

$$C_I = 2 \times \sum_{m=1}^{30} \sum_{i=1}^m \left(\sum_{j=i+1}^{30} y_jq_{ij} - \sum_{j=1}^i y_iq_{ij} \right) \times \left(\frac{d_{m+1}}{V} + t_d \right) y_{m+1} \times u_I, \tag{4}$$

Fig. 2 User benefits



where d_{m+1} represents the station spacing between station $m + 1$ and m , V is the transit speed, and t_d is the lost time at each station. The factor t_d accounts for the time lost through deceleration and acceleration as well as for dwell time at a station. No out-of-pocket costs are included in the user cost. Transit fares are not part of the user cost since they are merely transfer payments from users to operators. Thus, the user cost is equal to the waiting cost plus in-vehicle cost:

$$C_U = C_W + C_I. \tag{5}$$

The supplier cost (C_S) consists of three components as shown in Eq. (6):

$$C_S = C_C + C_O + C_M. \tag{6}$$

These are capital cost (C_C), operating cost (C_O), and maintenance cost (C_M). Capital cost (C_C) includes land acquisition, design, and construction, and rail-track laying costs:

$$C_C = \sum_t^{30} \sum_i^{30} (y_i^{(t)} - y_i^{(t-1)}) k_i, \tag{7}$$

where k_i is the fixed cost for link i . We use $y_i^{(t)} - y_i^{(t-1)}$, since k_i is the cost which only counts the first time when $y_i^{(t)}$ changes from 0 to 1. We assume (in Assumption 7 above) that some economies occur if several stations (and their links) are built together. In our numerical examples, the construction cost savings are set at 3 % for 2 stations, 6 % for 3 stations, 9 % for 4 stations, 12 % for 5 stations, 15 % for 6 stations, 18 % for 7 stations, 21 % for 8 stations, and 24 % for more than 9 stations.

The operating cost is the transit fleet size F_T multiplied by the hourly operating cost per car u_T (\$/vehicle-h) and the number of cars n_C needed per train. u_T includes the equivalent hourly capital cost of the rail cars. Because the optimal headway changes as we extend the line, we have to update the headway after every decision made. To obtain the fleet size, the transit round trip time $R^{(t)}$ is derived first as follows:

$$R^{(t)} = 2 \sum_i^{30} \left(\frac{d_{i+1}}{V} + t_d \right) y_{i+1}^{(t)}, \tag{8}$$

where d_{i+1} represents the station spacing between stations $i + 1$ and i . Since our demand function is not elastic with

respect to headway (which means that demand is fixed during each iteration), the optimal headway h can be found by checking the first-order derivative of the total cost (C) function with respect to h equal to zero and solving it for h . The second derivative of the total cost function with respect to h is also checked to insure that the total cost function is convex.

$$\frac{\partial C}{\partial h} = 0 \tag{9}$$

$$\frac{\partial^2 C}{\partial h^2} = \frac{2Rn_C u_T}{h^3} > 0 \tag{10}$$

The resulting optimal headway is

$$h^{(t)} = \sqrt{\frac{2n_C u_T y_{i+1}^{(t)} \sum_i^{30} \left(\frac{d_{i+1}}{V} + t_d \right)}{u_W \sum_i^{30} y_i^{(t)} q_{ij}^{(t)}}}, \tag{11}$$

where $h^{(t)} > h_{\min}$ (i.e., 0.0222 h) and $h^{(t)} < h_{\max}$ = train capacity/peak point one-way passenger flow.

The fleet size $F^{(t)}$ is then the transit round trip time divided by the headway h :

$$F^{(t)} = \frac{R^{(t)}}{h^{(t)}}. \tag{12}$$

With the fleet size, we can then determine the operating cost:

$$C_O^{(t)} = F^{(t)} n_C u_T. \tag{13}$$

Maintenance cost, C_M , is expressed as the passenger miles traveled (PMT) multiplied by a unit maintenance cost, u_L (\$/pass. mile):

$$C_M = 2 \times \sum_{m=1}^{30} \sum_{i=1}^m \left(\sum_{j=i+1}^{30} y_j q_{ij} - \sum_{j=1}^i y_i q_{ij} \right) \times \left(\frac{d_{m+1}}{V} + t_d \right) y_{m+1} \times u_L. \tag{14}$$

Therefore, the supplier cost is equal to the operating cost plus maintenance cost:

$$C_S = C_O + C_M + C_C. \quad (15)$$

Equations (16) to (21) present a model for maximizing the system's net present worth:

$$\text{Maximize NPW} = \sum_t^{30} (B - C)(1 + s)^{-t} \quad (16)$$

$$\text{Subject to } y_i^{(t)} = 1 \text{ or } 0 \quad (17)$$

$$y_i^{(t)} - y_i^{(t-1)} \geq 0, \quad \text{for all } i, t \geq 1 \quad (18)$$

$$y_i^{(t)} - y_{i+1}^{(t)} \geq 0, \quad \text{for all } t, i \geq 1 \quad (19)$$

$$f \times \text{revenue}^{(t)} \geq C_O^{(t)} + C_M^{(t)}, \quad \text{for all } t \quad (20)$$

$$(1 - f) \times \text{revenue}^{(t-1)} + \text{Subsidy}^{(t)} \geq C_C^{(t)}, \quad \text{for all } t. \quad (21)$$

Equation (16) is the objective function that maximizes the system's NPW. The annual net benefit is equal to total benefit (B) minus total cost (C). Total benefit includes user benefit; total cost includes supplier cost and user cost. We have to include the interest rate in the model to obtain the NPW. In Eq. (17) the decision variables are binary. Equation (18) is the realistic constraint ensuring that after link i is built, it always remains in operation. Equation (19) is the precedence constraint that prevents any link i from being built if any one of its predecessors is not yet completed. The transit line has to be built sequentially, since there would be fewer benefits if we randomly choose any segment to build along the route. In transit operation, some fraction of the fare collection may be used for covering operation expenses, and the remaining fraction (if any) may be used for funding the construction of new transit line extensions. Equation (20) is the revenue constraint for covering operational expenses, i.e., operating and maintenance costs. Due to uncertainties about the future, transit operators may try to balance their operation-related expenditures in each year. Thus, a fraction f of the revenue collected from fares is used to cover the operating and maintenance costs in each year. Equation (21) is the budget constraint for funding the capital investments. It shows that the construction costs in any year must not exceed the available capital funds plus some fraction $(1 - f)$ of the fare revenues accumulated from the previous year.

2.3 Simulated Annealing

Simulated Annealing (SA) is a heuristic method, which is very useful in optimizing objective functions with numerous local optima. It was originally developed by Metropolis et al. [15] who describe its details. Unlike most of the earlier search methods, SA may accept (with a decreasing

probability) moves to neighboring solutions which worsen the objective function, in order to escape from locally optimal "holes." Using SA, if a neighborhood solution is better than the previous one, it is always accepted. To avoid getting stuck in a local minimum or maximum, occasionally solutions worse than the current one are also accepted but with a probability similar to that in the dynamics of the annealing process. As the temperature decreases, the probability of accepting a bad solution is decreased and in the final stages the Simulated Annealing algorithm becomes similar to gradient-based search.

The simulated annealing process proceeds as follows:

Step 1 randomly generate a feasible initial solution x_0 and calculate $f(x_0)$.

Step 2 from the current solution x_0 , jump to its neighbor x' and calculate $f(x')$.

Step 3 compare $f(x_0)$ and $f(x')$.

If $f(x') > f(x_0)$, x' replaces x_0 to be the current solution.

Otherwise, randomly generate a number z between 0.01 and 0.99.

If $z < \exp\left(-\frac{f(x)-f(x')}{T}\right)$ [8], x' becomes the current solution.

Otherwise, do nothing.

Step 4 for every 5 iterations, reduce the temperature T by 1 %, i.e., multiplying by 0.99.

Step 5 check termination rule.

Maximum iterations reached or stopping criteria reached.

If yes, algorithm stops; otherwise, return to Step 2.

More detailed SA design and parameter tuning can be found in Cheng [9].

2.4 Numerical Results

The procedure was coded with MATLAB 7.2.0 and run on an IBM Laptop with a 1.60 GHz Pentium R processor and 1.00 Gigabytes of RAM. Since running a 30-station route over a 30-year analysis period takes considerable time, a very large number of iterations are needed to converge while searching with Simulated Annealing. In the numerical examples presented here, it is assumed for simplicity that the externally funded budget for capital improvements is equally distributed over all periods. Two problem cases were tested: an unconstrained case and a revenue-budget-constrained case.

2.5 Unconstrained Case

Figure 3a shows the resulting discounted net benefits/year and the optimized phases. Surprisingly, this optimized solution has only one phase which consists of adding 23 links

in year 2. Since it is assumed that we have unlimited funds for extensions, this answer implies that we should add links as soon as possible if the demand is sufficient. The demand at stations 28, 29, and 30 is initially too low, so the route stops at station 27. The annual discounted net benefits respond to the addition of links. In year 2, the negative value is due to the construction costs. Figure 4a compares four alternatives. The green line is the optimized solution found for the unconstrained case. The black line is the case without addition of links, which has only 4 stations in service for the 30-year horizon. The drop in year 2 is due to capital costs for extension. If the transit line is extended to link 27 in year 2, the NPW will increase much faster than without an extension. Alternative 1 (red) extends to link 27 in year 17; alternative 2 (blue) extends to link 30 in year 2. None of them have a higher objective value than the green line.

2.6 Constrained Case

Two kinds of constraints are added: a revenue constraint and a budget constraint. Penalty methods are used here for dealing with constraints. A 5 % borrowing allowance is added into both revenue and budget constraints. Adding such an offset is reasonable to avoid delaying the construction just because of small shortfalls.

For the revenue-budget-constrained case, the stopping criterion is increased to 100 k iterations and the objective value is 4.0591×10^9 . Figure 3b shows the annual discounted net benefits in each year and the optimized phases. There are six phases for this case: Phase I adds 3 links in year 3; Phase II adds 2 links in year 5; Phase III adds 1 link in year 6; Phase IV adds 1 link in year 9; Phase V adds 3 links in year 13; and the last phase adds 1 link in year 14. The annual discounted net benefits drop significantly when links are added but bounce back with a higher value the following year. Figure 4b shows the NPW for different cases. In Fig. 4b, as more constraints are applied, NPW decreases, as expected. However, the differences in NPW between revenue-constrained case and revenue-budget-constrained case are small. There are probably two reasons: first, the 5 % borrowing allowance brings the answers in our two cases fairly close; second, the revenue constraint dominates in the numerical example. Adding a budget constraint does not bind the solution. Compared with the unconstrained and revenue-budget-constrained cases, the NPW in the case constrained by revenue and budget is nearly one-third of that in the unconstrained case. NPW is significantly affected by the construction phases.

2.7 Reliability

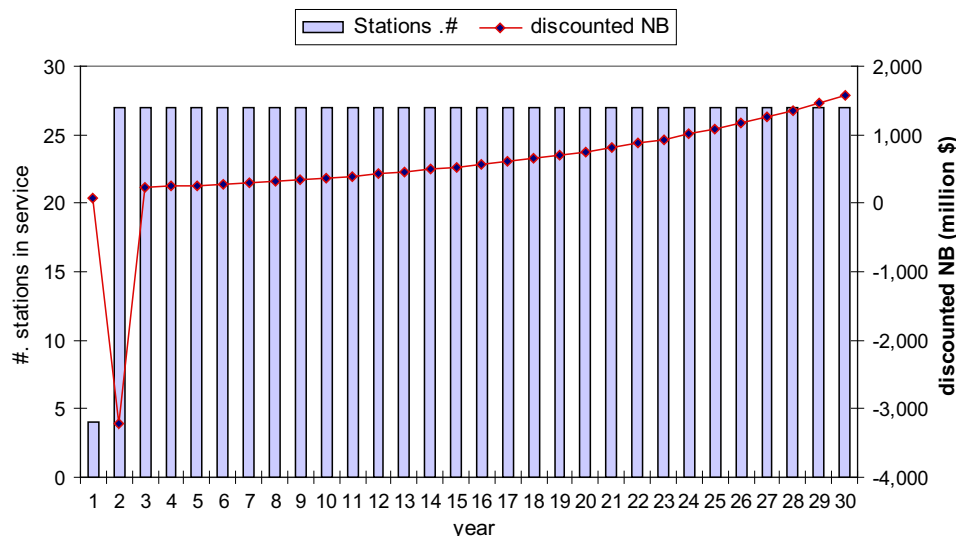
The reliability of the obtained solution is an important concern. Since the exact optimal solution to this problem is

not known (note that no existing methods guarantee finding the global optimum for a large RCPS), it is difficult to prove the goodness of the solution found by the proposed Simulated Annealing algorithm. Therefore, an experiment is designed to statistically test the effectiveness of the algorithm. In this experiment, the fitness value is evaluated for each randomly generated solution to the problem. First, numerous solution samples are generated and tested. The next step compares the random sample solutions with the SA optimized solution.

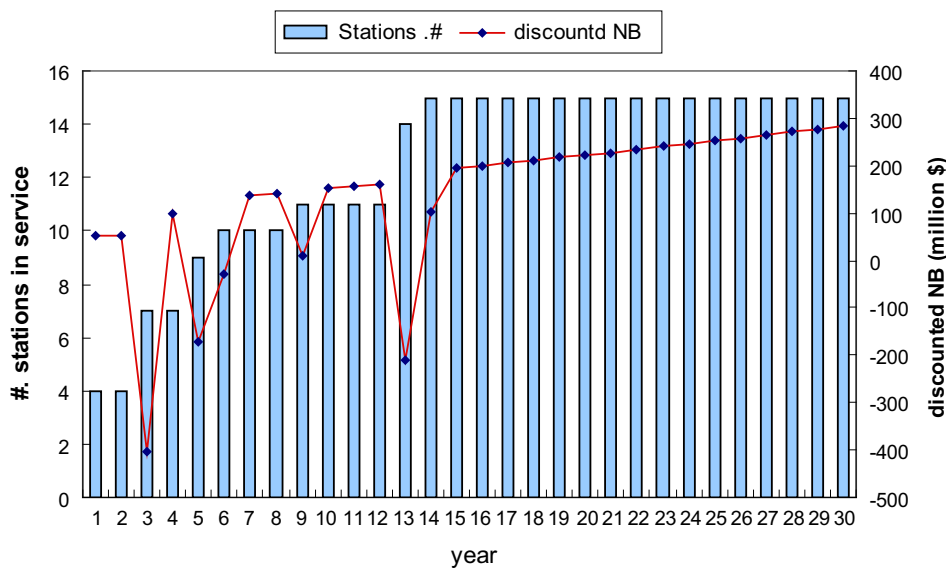
We first create a random sample of 1,000,000 observations. The best fitness value (i.e., NPW) in this sample is 3.7527×10^9 , while the worst one is -1.0278×10^{12} . The sample mean is -2.2814×10^{11} and the standard deviation is 1.4481×10^{11} , as shown in Fig. 5. In the experiment, the optimized solution obtained (4.0591×10^9) is approximately 8 % higher (i.e., better in NPW) than the highest value in the random sample (3.7527×10^9). In other words, the solution found by the SA algorithm dominates by a considerable margin all the solutions in the distribution. In fact, the random sample does not cover the range of the fitness values for all possible solutions in the search space. The number of possible solutions for the unconstrained case is 27^{29} , which includes infeasible solutions. This number comes from the solution vector which has 30 elements. Besides the base year (year 1), in each year the number of stations in service can change from 4 to 30, so there are 27^{29} different permutations. It is difficult to calculate the exact number of possible feasible solutions, since the problem is dynamic. This suggests that an even larger sample might be worth testing. However, the optimized solution value is considerably better than any of the 1 million random solutions sampled. The result shows that the best solution found by the SA algorithm, although not necessarily globally optimal, is still remarkably good when compared with other possible solutions in the search space and is unlikely to be significantly improved upon by the globally optimal solution. We can conclude that the solution quality will be limited by the various uncertainties regarding the inputs rather than the capabilities of the Simulated Annealing algorithm. This analysis indicates a very promising performance for the proposed optimization model.

2.8 Computation Time

One of the main drawbacks of the Simulated Annealing approach is its computation time. As the problem size changes from ten stations and 10 years to thirty stations and 30 years, respectively, the computation time increases significantly, as shown in Fig. 6. Various computations such as computation of the net present worth function and computation of the probability of accepting bad solutions



(a) Unconstrained Case



(b) Constrained Case

Fig. 3 Discounted net benefits and optimized phases

increase the computation time when the problem size grows. Also, for better results the cooling schedule has to be carried out very slowly and this significantly increases the computation time.

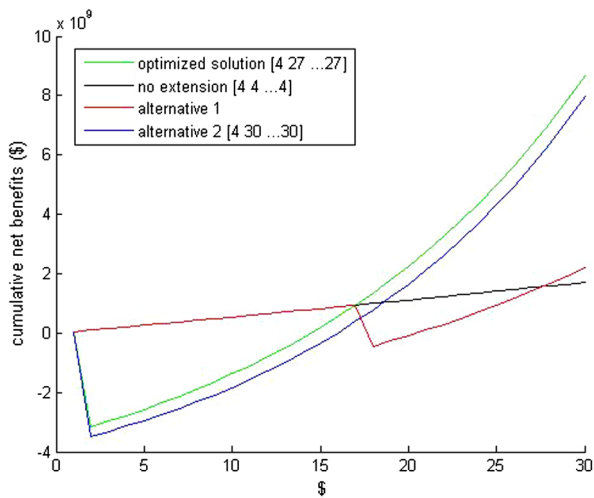
2.9 Sensitivity Analysis

The following sensitivity analysis is designed to investigate the effects of one input parameter (i.e., the interest rate) on the resulting optimized values (i.e., construction phases and total net benefits). If the model is very sensitive to changes in a particular input parameter, that parameter should be predicted as accurately as possible and decisions should be made more cautiously.

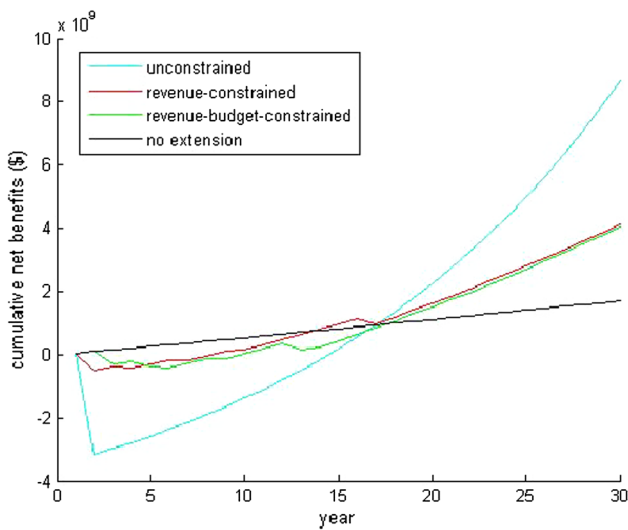
2.9.1 Interest Rate

The interest rate plays an important role in project scheduling, especially in large investment projects. Theoretically, projects tend to be postponed when the interest rate is high. If the interest rate increases, then investment decreases due to the higher cost of borrowing. Although transit planners cannot control the interest rate, sensitivity analysis can show them how extension decisions are affected by interest rates.

To evaluate the effects of different interest rates (s) on phasing decisions and NPW in this section, s , whose base value is 5 %, is varied between 0 and 30 %. Table 2 shows the differences in optimized values and phases. Not only is



(a) Comparison of Alternatives for the Unconstrained Case



(b) Comparison of All Cases

Fig. 4 Cumulative net benefits over years on different alternatives and cases

the extension postponed but also the number of phases decreases when the interest rate increases. When s is below 10 %, the transit line is extended to link 15. When s increases to 15 %, the line is extended to link 8. When s exceeds 30 %, the transit route merely extends to link 5. For links with enough demand, delaying the construction causes no problem. The marginal benefits of adding links with enough demand are always positive, except when adding links in the last year of the analysis period. However, links with initially low demand and enough high growth rates after implementation are only beneficial over the analysis period if they are built early. In order to achieve higher cumulative net benefits, the links which would be economically beneficial at the end of the analysis period must be added as soon as possible. If some

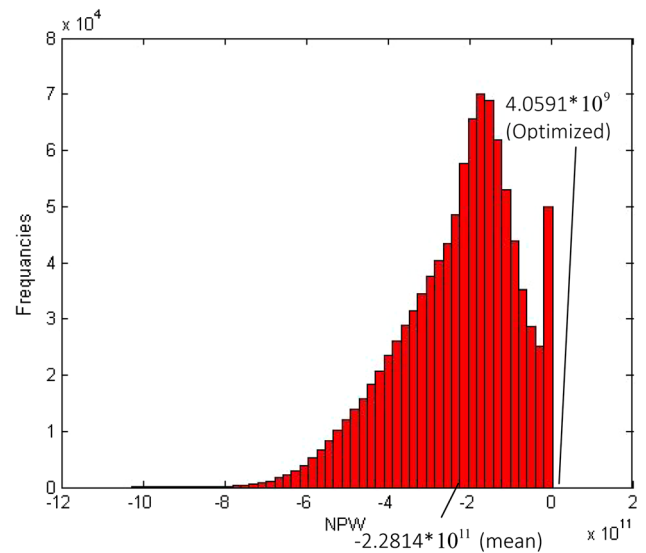


Fig. 5 Optimized SA solution compared to 10^6 random solutions

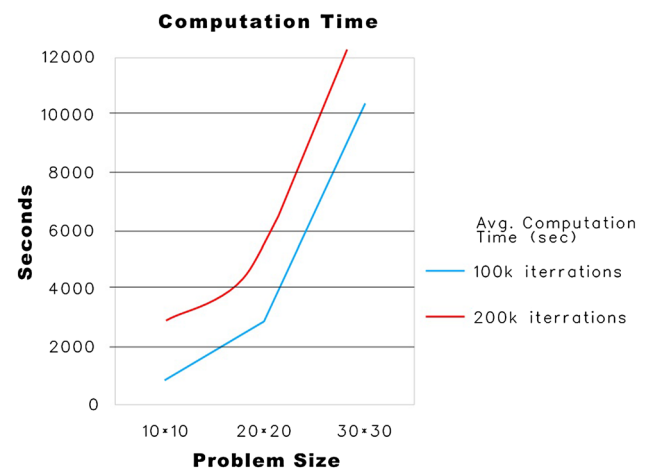


Fig. 6 Computation time

constraints prevent the extensions at early stages, the line cannot be extended as far as in the unconstrained case.

3 Conclusions

A model is developed for optimizing the construction phases of any rail transit line that is built without gaps from one end toward the other. It can be used to determine not only the construction phases but also the economic feasibility of additional links under various financial constraints. The optimized solution also avoids overextension of the proposed line. In addition, tax-funding policy also can be optimized through sensitivity analysis, as demonstrated.

The study leads to the following conclusions:

The numerical analyses show that for the unconstrained case, immediately adding all links with positive net

Table 2 Effects of interest rates on NPW and optimized phases

Interest rate (%)	Cumulative net benefits (\$)	# phases	# Stations in service in each year																														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1.20E + 10	8	4	5	7	7	9	10	10	10	11	11	13	13	14	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
3	6.19E + 09	8	4	5	7	7	9	10	10	10	11	11	13	13	14	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
5	4.05E + 09	6	4	4	7	7	9	10	10	10	11	11	11	11	14	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
7	2.70E + 09	6	4	5	7	7	9	10	10	10	11	11	11	11	11	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
10	1.53E + 09	4	4	4	4	8	9	9	9	9	9	9	9	9	9	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
15	7.07E + 08	2	4	4	4	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
20	4.12E + 08	3	4	4	4	5	5	5	5	5	5	5	5	5	5	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
30	2.39E + 08	2	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

benefits achieves the highest objective value. This result is consistent with the one found in Kolisch and Padman [9], which is to schedule jobs with positive cash flows as soon as possible and to delay jobs with negative cash flows as much as possible. With its given inputs, the optimized solution has only one phase and, in the absence of a completion constraint, does not reach the end of the route. Therefore, those links with negative values are postponed indefinitely. If we insist (through completion constraints) that outer links with unjustifiably low demand must be completed, then those links with insufficient demand would be added in the last time period (The present worth of their costs would thus be minimized). For the case in which demand grows faster after an extension, the economic feasibility of adding one link is affected significantly by the construction time. Compared with various financial constraints, the transit line can be extended to link 27 for the unconstrained case, but it can only be extended to link 15 for the case constrained by external budget and route-generated revenue. If some links with low demand and high growth rate after extension cannot be added at early stages, they do not become justified within the remaining 30-year span of our case study. That is due to the high capital costs of adding links. In our sensitivity analysis, no extension was justified at later stages. Consequently, when analyzing the economic feasibility of a project with high capital cost, construction phases should be taken into account.

The results obtained are reasonable, even for the possibly counterintuitive results where demand growth accelerates after links and stations are added. Such a model is valuable because it quantifies the effects of extension alternatives and finds extremely good solutions for this large combinatorial problem. While most of the results seem reasonable or even obvious qualitatively, such a model can help quantify and optimize the route development decisions.

3.1 Future Research

The following extensions are suggested for further studies:

- (1) The model designed in this study is deterministic. Based on uncertainties about the future, this model could be improved to consider probabilistic factors. For instance, the demand growth rate might change over time. Demand will not necessarily increase in the future. Interest rates and inflation rates also vary over time. A probabilistic model can address this problem more realistically than a deterministic model.
- (2) For increased realism, a future model might relax some simplifying assumptions, such as that specifying

sequential link addition. Currently the model can be used for radial networks. For some other cases, the assumption that only adds links sequentially should be relaxed.

- (3) Additional factors that would be affected by transit extensions might be modeled, such as multi-modal access to stations. External benefits and costs can be added into the model if they are correctly estimated, including employment opportunities, land values, travel time savings, and environmental impacts.
- (4) Some operational variables (e.g., transit fare and cruise speed) can also be optimized by a modified model at various times instead of keeping them fixed. In order to optimize these variables, price and travel time elasticity of the demand would have to be considered.
- (5) This model optimizes the construction phases for a single route. It might be improved to deal with more complex networks that include branched routes.
- (6) Other metaheuristic algorithms, such as genetic algorithms and tabu search, might be tried for this problem in attempting to reduce the computation time.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Chien S, Schonfeld P (1998) Joint optimization of a rail transit line and its feeder bus system. *J Adv Transp* 32(3):253–284
2. DiJoseph P, Chien S (2013) Optimizing sustainable feeder bus operation considering realistic networks and heterogeneous demand. *J Adv Transp* 47(5):483–497
3. Fan W, Machemehl RB (2006) Optimal transit route network design problem with variable transit demand: genetic algorithm approach. *J Transp Eng* 132(1):40–51
4. Farahani RZ, Miandoabchi E, Szeto WY, Rashidi H (2013) A review of urban transportation network design problems. *Eur J Oper Res* 229(2):281–302
5. Guan JF, Yang H, Wirasinghe S (2006) Simultaneous optimization of transit line configuration and passenger line assignment. *Transp Res Part B* 40(10):885–902
6. Guihaire V, Hao JK (2008) Transit network design and scheduling: a global review. *Transp Res Part A* 42(10):1251–1273
7. Kim M, Schonfeld P, Kim E (2013) Comparison of vertical alignments for rail transit. *J Transp Eng ASCE* 139(2):230–238
8. Kim M, Schonfeld P (2014) Integration of conventional and flexible bus services with timed transfers. *Transp Res Part B* 68B–2:76–97
9. Kolisch R, Padman R (2001) An integrated survey of deterministic project scheduling. *OMEGA Int J Manag Sci* 29:249–272
10. Lai X, Schonfeld P (2012) Optimization of rail transit alignments considering vehicle dynamics. *Transp Res Record* 2275:77–87
11. Li ZC, Lam WHK, Wong SC, Sumalee A (2012) Design of a rail transit line for profit maximization in a linear transportation corridor. *Transp Res Part E* 48(1):50–70
12. Lo HK, Szeto WY (2009) Time-dependent transport network design under cost-recovery. *Transp Res Part B* 43(1):142–158
13. Markovic N, Milinkovic S, Tikhonov KS, Schonfeld P (2015) Analyzing passenger train arrival delays with support vector regression. *Transp Res Part C* 58:251–262
14. Matisziw TC, Murray AT, Kim C (2006) Strategic route extension in transit networks. *Eur J Oper Res* 171:661–673
15. Metropolis NA, Rosenbluth A, Rosenbluth M, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092
16. Shayanfar E, Abianeh AS, Schonfeld P, Zhang L (2015) Prioritizing interrelated road projects using meta-heuristics. *J Infrastruct Syst ASCE*
17. Spasovic LN, Schonfeld P (1993) Method for optimizing transit service coverage. *Transp Res Rec* 1402:28–39
18. Sun Y, Schonfeld P (2015) Stochastic capacity expansion models for airport facilities. *Transp Res Part B* 80:1–18
19. Szeto WY, Jaber X, O'Mahony M (2010) Time-dependent discrete network design frameworks considering land use. *Computer-Aided Civil Infrastruct Eng* 25(6):411–426
20. Tsai FM, Chien S, Wei CH (2013) Joint optimization of temporal headway and differential fare for transit systems considering heterogeneous demand elasticity. *J Transp Eng ASCE* 139(1):30–39
21. Valadares Tavares L (1987) Optimal resource profiles for program scheduling. *Eur J Oper Res* 29:83–90
22. Vuchic V (2005) *Urban transit: operations, planning and economics*. John Wiley & Sons Inc, Hoboken, NJ
23. Wang S, Schonfeld P (2007) Demand elasticity and benefit measurement in a waterway simulation model. *TRB 86th annual meeting compendium of papers*
24. Zhou Y, Kim HS, Schonfeld P, Kim E (2008) Subsidies and welfare maximization tradeoffs in bus transit systems. *Ann Reg Sci* 42–3:643–660