Computer-Aided Civil and Infrastructure Engineering **24** (2008) 1–14

Optimization of Large Transport Networks Using the Ant Colony Heuristic

Basil J. Vitins*

Rapp Trans AG, Hochstr. 100, 4018 Basel, Switzerland

&

Kay W. Axhausen

Institute for Transport Planning and Systems, ETH Zurich, 8093 Zurich, Switzerland

Abstract: *Long-term transportation planning for larger regions must assess synergies and interference among sets of projects. The objective is the maximization of the overall benefit within specific budget restrictions by finding the most favorable bundle of projects, that is, solving the network design problem. For large numbers of projects, complete enumeration of all combinations is not feasible for detailed networks. The ant colony heuristic is suitable for this kind of problem. According to our knowledge, this article presents the above-mentioned heuristic's first application to a realistically sized network. A detailed multimodal network assignment of a substantial Swiss city provides the basis for calculations. First, each infrastructure project is assessed using a cost–benefit analysis. The ant colony heuristic is then successfully executed and the bundles are evaluated. The article provides new insights into applications of the heuristic in large networks and focuses on problematic calibration details as well as the choice of objective function. Suggestions are made for general applications and further research.*

1 INTRODUCTION

Congested roads and overcrowded transit lines increase travel time and costs, and lead to economic, social, and environmental losses. Today, no one questions the need to expand numerous urban transportation networks. New network infrastructure can result in large societal benefits. The existing network can either be improved through improvement of existing roads and transit lines or by building additional links and lines to extend the network. In master plans, such expansion can be combined with the removal of existing infrastructures because of spatial, environmental, and societal considerations. Network improvement must be carefully considered, especially within heavily used transportation networks. Negative consequences of new infrastructure are often very difficult to predict, so the actual benefit–cost ratios could be smaller than expected.

When evaluating new infrastructure projects, various qualitative and quantitative assessment methodologies can be employed, for example, Multi-Attribute Utility Theory (Keeney and Raiffa, 1976) or cost–benefit calculation (Stopher and Meyburg, 1976). Most of existing methodologies focus on already defined projects or project bundles. However, in transportation networks, the best bundle of projects is usually not obvious. Often, one has to choose the best alternative from a very large number of possible bundles. In contrast to other policy areas, evaluations of network change have to consider interdependences between projects. Therefore, when implementing two projects, assessment outcomes cannot be summed up to obtain overall benefit. Instead,

*To whom correspondence should be addressed. E-mail: *basil.vitins@rapp.ch*.

the assessment outcome must be recalculated to include both projects simultaneously. The benefit–cost ratio of the joint scenario will most likely be different from the average cost–benefit ratio of the independent projects. Moreover, additional links do not automatically increase the overall benefit (Braess, 1968). It has been shown that new links can also increase travel time, even though an improvement was expected, for example, when building a shortcut. The well-known knapsack problem and its methodologies could be applied when assuming independence (or only limited interactions) between the projects (Kellerer et al., 2004). As this assumption is unrealistic, it would be very useful to have a procedure or algorithm that can address these interactions appropriately. This article will present and apply a suitable heuristic algorithm to address this network design problem (NDP), that is, the selection of the optimal bundle of discrete projects from a larger set, while accounting for the interaction of their effects on network flows.

Although other papers focus more on methodology of the NDP solution, this article aims to provide a link between theory and practice. We transfer the approach in Section 2, suggested by Poorzahedy and Abulghasemi (2005), and apply it to a large bimodal transportation network (350 zones) of the Bern region (Switzerland). Realistic infrastructure projects replicate the real-life situation as far as possible. The region encompasses highly congested streets, occasional parallel routes and shortcuts. Complex route choice and secondary effects call for careful, detailed modelling and planning if the current situation is to be improved efficiently.

First, the article formally defines the NDP (Section 2). A short overview summarizes existing methodologies, and the next section explains the methodology applied in detail. Then, the study area and its characteristics are introduced to the reader (Section 3). Special attention will be given to the two mode choice models, used alternatively. Results on the behavior of the solution algorithm (particularly its convergence) are presented, and finally, the bundles obtained for the study region are discussed; suggestions for further work conclude the article (Section 4).

2 NETWORK DESIGN PROBLEM

The NDP has recently been revisited by various authors, both because computation speeds have increased, and because the issue has become more important, especially in regard to planning for emergencies, evacuations, and reliability improvement (Bakuli and Smith, 1996; Bell, 2000; Bonabeau et al., 2000; Sumalee et al., 2006; Tuydes and Ziliaskopoulos, 2006). This section will briefly review the problem and suggested solutions.

2.1 Definition

NDP is a discrete optimization problem and NP-hard (Garay and Johnson, 1979). It was first defined by LeBlanc (1975). The group of NDP includes discrete, continuous, and mixed NDP. The discrete NDP only refers to fixed projects, unlike to the continuous NDP, which includes projects with continuous costs and capacities. Mixed NDP is a combination of the first and second problem. For practical reasons and simplicity, only the discrete NDP is considered here; in addition, the selection among discrete projects is a typical real-world application. The NDP can be defined in combination with various combinatorial problems. Here, the NDP is defined for a static demand transportation network with links, corresponding volume delay functions, and nodes, as it is used in many cities for planning purposes. Additional modes, functions, and parameters normally used in a transportation model can be included as well. The goal of the NDP is to choose a bundle among a defined set $x_1 \dots x_n$ of possible infrastructure projects i; while $x_i = 0, 1$ and $x_i = 1$ when project i belongs to the bundle. Moreover, x_i belongs to vector $\mathbf{x}$ and means that the project i will eventually be implemented in the network with all other projects j, for which x_j are set to 1 as well. Projects can include public or private transportation changes, for example, a new tramway or a new bypass. When $x_i = 0$, the project i will be excluded from the bundle. The chosen project bundle has to generate maximum benefit $\mathbf{c}$, whereas the benefit of project i is a function f_i of the other projects chosen. The calculation of the benefit N, using f_i, is crucial and includes here the time-consuming equilibrium assignment problem of the network simulation. Normally, the benefit of all chosen projects is calculated in combination to satisfy the goal function $\boldsymbol{c}^T\boldsymbol{x}$, which means the maximization of the "overall" benefit. Here, we use cost–benefit analysis to calculate f_i (see below for details). During cost–benefit analysis, variables are normally compared with a reference scenario that contains no changes in demand and network. Constraints limit the costs a_i of all projects i between zero and the budget constraint A. The summed costs $\sum a_i \cdot x_i$ have to be below the predefined budget constraint A.

$$
\begin{aligned}
\text{maximize} \quad & \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & \mathbf{a}^T\mathbf{x} \leq A \\
& c_i = f_i(\mathbf{x}) \\
& \mathbf{x} \in \{0, 1\}^n
\end{aligned}
$$

$$
\text{where as} \quad \mathbf{a} > \mathbf{0};\ A > 0;\ \max_{j=1,\dots,n} a_j \leq A < \sum_{j=1}^{n} a_j
$$

2.2 Existing methodologies to solve the NDP

A complete enumeration of all possible combinations is the only way to determine the correct and exact solution of the NDP. In our case, the estimated computation time for a complete enumeration of all possible bundles of 14 preselected projects is 114 days, and is not normally practical. All other methodologies are either heuristics or based on relaxations of certain constraints. The well-known knapsack approach and its corresponding solutions (Kellerer et al., 2004) can be applied when ignoring possible interactions between projects. The quadratic knapsack problems and its existing solutions (Kellerer et al., 2004) can be applied if there are only pairwise interactions between projects. Hsieh and Liu (2004), Taber et al. (1999), and to a certain extent Sumalee (2007) propose genetic algorithms to solve the discrete NDP and similar problems. Regression analysis is another, less complicated way to calculate a possible solution. Optima Consortium (1997) proposes a step-by-step regression analysis to find the optimal bundle.

The ant colony heuristic, used here, mirrors ant behavior, a form of artificial or swarm intelligence (Bonabeau et al., 2000). This methodology has been employed for a variety of operational research problems and is currently known as the best algorithm for many of them (Bonabeau et al., 2000; Dorigo et al., 1999; Merkle et al., 2002). The ant colony has a straightforward structure and can be implemented in different programming environments requiring only basic programming knowledge. Regarding computation times, and problem setup, the ant colony heuristic belongs to one of the most competitive methods (Bonabeau et al., 2000). The chosen heuristic is also especially suited for discrete problems. Alternative methodologies like the knapsack problem and its extensions do not satisfy the preconditions of the NDP because the necessary relaxations violate constraints of the NDP. The ant colony algorithm belongs to the family of evolutionary algorithms, which are very promising regarding very large data sets as well as NP-hard problems. Genetic algorithms are a legitimate alternative to the ant colony heuristic and have been already applied in different cases. However, the complex structure requires an additional effort in implementation and the algorithm may be less practicable in application.

2.3 Ant colony heuristic

The ant colony heuristic can be explained by an analogy with ants' social behavior. Ants communicate through chemical substances, called pheromones, while looking for food resources. Ants notice pheromones deposited by previous ants on the trails. The more a trail is used, the more pheromones are dispensed on it, allowing ants to learn from the success or failure of preceding ants. As opposed to ant behavior, pheromones in the algorithm are only deposited when the entire tour is finished and the overall success is known.

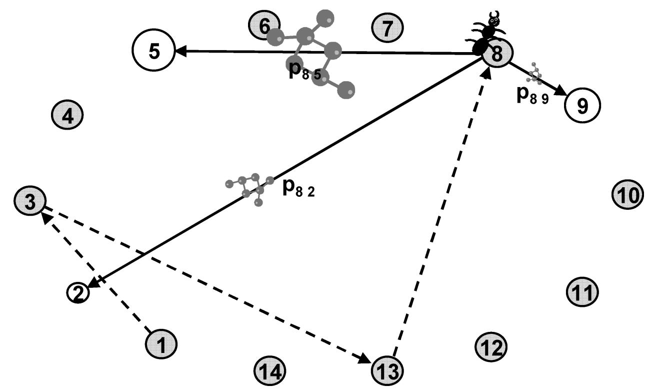

Fig. 1. The decision-making process of an ant.

Figure 1 shows different network optimization projects $i = 1 \ldots 14$ schematically as circles arranged as an ellipse. The project or circles are numbered from 1 to 14, whereas the actual sequence of the projects is irrelevant. In Figure 1, the ant has already chosen project 1, 3, 13, and 8 and is now continuing onto one of the projects $j \in \{2, 5, 9\}$.

The lastest project chosen is called project i. The ants choose additional projects j according to the probability p_{ij}, which depends on the current pheromone density τ_{ij} on the corresponding link $[i \;\; j]$ and the benefits N_{ij} of the projects j (P1). Benefit N_{ij} as it is mentioned above is calculated according to Stopher and Meyburg (1976); see Section 2.4 for more details on the calculation of benefit N_{ij}.

$$p_{ij} = \begin{cases} \dfrac{e^{\alpha\tau_{ij}} \cdot e^{\beta N_{ij}}}{\sum\limits_{m \in F_k} e^{\alpha\tau_{im}} e^{\beta N_{im}}}, & \text{when } j \in F_k \\ 0, & \text{otherwise} \end{cases} \qquad \text{P1}$$

The approach P1 above was formulated in Dorigo et al. (1996) and Poorzahedy and Abulghasemi (2005) and has a multinomial logit structure. The simple structure allows clear and comprehensive parameter definition and a comprehensive calibration. Projects with large benefits (when assessed separately) and high pheromone density on link $[i \;\; j]$ are chosen more often. In an extended approach, the benefit of project j would account for the interaction with project i as well. Here, for a reduced computational burden, N only includes the benefit of project j ($N_{ij} = N_j$). So a small variation of p_{ij} is accepted intentionally to lower calculation time considerably but without affecting the calculation of the overall benefit. F_k is the set of available projects from which the ant chooses one particular project according to p_{ij}, and depends on the financial resources left. Projects in dark circles in Figure 1 do not belong to F_k and therefore cannot be chosen anymore: $F_k = \{2, 5, 9\}$. In addition, α and β are parameters requiring calibration.

When $F_k = \varnothing$, meaning that the financial resources are depleted, the projects are implemented simultaneously in the transportation network, and sequence does not matter anymore. The overall benefit of the selected bundle is calculated using f_i (see Section 2.4 for details) and compared with previous bundles that are calculated before in the same manner. So an iterative process occurs whereas the ant chooses projects and creates new bundles for assessment. Following Poorzahedy and Abulghasemi (2005), one iteration is defined as 14 tours of different ants, leading to 14 project bundles and 14 different overall benefit calculations. Supplementary, the algorithms report better results when ants start at each project once, instead of starting at a randomly chosen project.

The pheromone amount, which is distributed on the links, depends on the benefit of the entire bundle created. The larger the overall benefits, the larger the amount dispensed on the links. So the algorithm accounts for possible interactions between different projects and negative benefit as well (Braess, 1968). In addition, to improve convergence, an evaporation rate is included in the algorithm (Poorzahedy and Abulghasemi, 2005). The pheromones eventually decay unless more ants follow onto the same links. Generally, accurate evaluation of the evaporation rate increases in importance with the difficulty of the optimization problem (Dorigo and Stützle, 2004). Furthermore, according to Poorzahedy and Abulghasemi (2005), the algorithm obtains better results when pheromone density is not recalculated after each ant tour. We recalculate pheromone density after 14 bundles, or one iteration. A fraction ρ of the existing pheromone molecules τ_{ij}^0 evaporates after each iteration, before the new molecules $\Delta\tau_{ij}^k$ are added to the links (P2).

$$\tau_{ij}^1 = \rho \cdot \tau_{ij}^0 + \sum_{k=1}^{m} \Delta\tau_{ij}^k \qquad \text{P2}$$

P2 follows Poorzahedy and Abulghasemi (2005) and has, for simplicity, a clear linear structure. A complex calibration process, as it is applied here, profits from a comprehensive evaporation formula like P2. In addition, ρ could simply be adjusted during calculation; this technique is suggested for further research. $\Delta\tau_{ij}^k$ is the benefit of bundle k, α (P1) is responsible for the influence of τ_{ij}^1. Link $[i\ \ j]$ was chosen m times before the new density is calculated. Pheromone densities are initialized to zero before starting the algorithm. Here, ρ is calibrated together with α and β for optimal convergence (see Results section).

In summary, the algorithm proceeds as follows:

Step 1: Benefit (N) is calculated for each project, using f_i. In this work, we use cost–benefit analysis to calculate N; other assessments are possible as well.

Step 2: The first ant starts at project 1 and chooses a second project according to p_{ij} (P1). As soon as the financial resources are depleted and no feasible projects exist anymore ($F_k = \phi$, whereas F_k is the set of feasible projects), the overall benefit–cost ratio N is calculated for the evaluated bundle.

Step 3: Step 2 is repeated until a certain number of bundles k ($k = 14$ in this work) are defined by the ants. The ants start at each project once, instead of choosing the initial project randomly, as suggested by Poorzahedy and Abulghasemi (2005).

Step 4: Pheromone density is calculated for each link according to formula P2.

Step 5: As soon as the benefit–cost ratio of the best bundle (of one iteration) does not increase anymore, pheromone densities on each link are slightly changed to avoid a local optimum. We accomplished this with the method suggested by Poorzahedy and Abulghasemi (2005): The amount of pheromones are doubled on links with pheromone density below average. Thus, the algorithm is able to leave the local optimum and move toward the global optimum.

Step 6: When the overall benefit does not change anymore after three iterations (Poorzahedy and Abulghasemi, 2005), the algorithm stops, otherwise it proceeds to Step 2.

2.4 Assessment of the bundles

A key advantage of the ant colony heuristic is its independence from the assessment methodology employed. Here, the isolated projects and the bundles are assessed using cost–benefit analysis (Stopher and Meyburg, 1976). Within the definition of the ant colony algorithm (Section 2.1), the assessment function is called f_i and is applied at the end of each ant tour, after implementing the projects in the network simulation. During assessments, variables are compared with a reference scenario with no demand and infrastructure changes. Cost–benefit analysis is a popular and widely accepted evaluation tool in Switzerland (VSS, 2006) and elsewhere. Other methodologies could also be applied with the ant colony heuristic. However, only methods not requiring the analyst's intervention during an iteration are appropriate. Here, the benefit–cost function is programmed in EMME, a multimodal transportation planning software, and the benefit–cost function is calculated automatically after each ant tour. For simplicity, only travel times and distances are considered in the benefit–cost analysis employed here, but account is taken of modal shifts due to network improvements. The equilibrium assignment problem has to be solved for public and private transportation to calculate travel times and distances. Additional indicators such as environmental impacts, safety gains, etc. could be included as well but had to be excluded owing to data availability problems; for details, see Stopher and Meyburg (1976) or VSS (2006). Travel time savings are valued according to Axhausen et al. (2008) and VSS (2007), which vary the valuations by distance travelled. Travel distances are used to calculate the operating costs of private transportation. Operating costs of public transportation are not included. No changes in trip generation and distribution are integrated into evaluation. For this reason, the assessment function calculates only the First Year Return (benefit–cost ratio of the first year after opening). As an extension, it is possible to account for demographic changes within demand calculations. Such calculations are left out for the sake of computation time and because the primary aim of this article is to implement the ant colony heuristic. A time horizon of 40 years is assumed for all calculations. Benefits and costs are discounted (VSS, 2006).

3 REGION OF BERN

The case study reported here is the first with a realistically sized network. The algorithm was implemented using EMME's macro language (INRO, 1998), interfaced with the existing static EMME transportation network of Bern (RVK4 et al., 2004), which is applied in current transportation planning tasks. Because of ongoing work for a new long-range transport master plan, the case study makes use of the slightly out-of-date model for the year 2000.

3.1 Study area

The perimeter of the model surrounds the capital Bern and encompasses 370,000 inhabitants; 235,000 employees (year 2002); and 85 municipalities. Owing to the fact that the city of Bern is the capital of Switzerland, there is a rather high ratio of employees to residents. The city of Bern has 129,000 inhabitants. The suburban areas and the city of Bern are highly congested during morning and evening peak hours. The main highways around the city are at their capacity limits. Compared to off-peak times, private transport travel times across the study area increase by 10% during rush hour due to congestion effects. There is a clear need to improve the current situation.

3.2 Multimodal simulation of the region of Bern

The transportation model used in this work has the characteristics listed in Table 1. Network and demand uses data of the year 1998 or later. For simplicity, only the evening rush hour (5 pm–6 pm) is considered. Demand includes trips in and out of the perimeter, through trips, and local trips. Demand for private and public transportation only changes through mode choice, whereas trip generation and spatial redistribution are neglected. For the assessment, travel time is calculated over the entire modelled study area.

Private transportation route choice depends on link travel times, which are functions of actual traffic flow. Route choice is calculated using the well-known Wardrop Equilibrium (Frank and Wolfe, 1956; Wardrop, 1952). In the public transportation assignment, passengers are assumed not to know timetables and to choose randomly out of a set of feasible line combinations (Spiess, 1984; Spiess and Florian, 1989). The probability of choosing a feasible line is proportional to the frequency. In-vehicle time, access time, boarding and alighting is weighted during the assignment to reproduce passenger cost perception. Public transportation equilibrium does not account for any congestion effects.

Table 1
Characteristics of the network analyzed

Number of zones	350
Modes	Public transportation (train, bus, and tramway) Private transportation
Nodes	724
Links	2,100
Demand public transportation (evening peak hour):	38,500 [pers.]
Demand private transportation (evening peak hour):	69,500 [veh.]
Travel time public transportation (evening peak hour):	58,500 [h∗per.]
Travel time private transportation (evening peak hour):	46,100 [h∗veh.]

3.3 Modal split estimation

Two different modal split parameter sets were tested to assess result sensitivity. The first set is derived from the neighboring Zurich area mode choice model (Vrtic et al., 2005). This modal split model considers travel time and distance, access time to public transportation, headway, and the number of transfers. Fares are not taken into account directly but they are included via the highly correlated travel distances. The parameters were estimated using a maximum-likelihood approach and are applied in the public transportation model of Zurich, the largest city of Switzerland. Owing to lack of time, the modal split parameter for the logit model are adopted from Zurich to have a travel behavior as close to Bern as possible. The second parameter set had been calibrated for Bern (RVK4 and RappTrans, 2002), and has a simpler general cost formulation compared with the Zurich model. It includes travel time and distances and distinguishes between suburban areas and the city center. Its modal split calculation is based on a multinomial logit model, just as in Zürich. It encompasses fewer parameters than the model of Zurich but was calibrated especially for the transport model of Bern used in this work. It is expected to be less sensitive to network and service changes.

3.4 Potential infrastructure projects

From the set of projects currently discussed in Bern, a set of 14 was selected. This set includes both small and large projects, and public transportation as well as road projects (Table 2 and Figure 2). Values in Table 2 are estimates and cannot be compared with ongoing assessments of similar projects. The budget constraint imposed for the algorithm is 3,000 Mio sFr. For lack of space, project details were left out; implementations in the network model are adequate and are consistent with the overall network simulation.

4 RESULTS

4.1 Modal split calculations

Aggregated travel time changes are central elements of any transport project benefit. In a congested urban environment, it is necessary to assess the intermodal competition as well. In advance of the optimization, the size of these interactions was assessed by calculating the impacts of modal shifts on public and private transport travel times using the two models described earlier, iterating once between assignment and mode choice. To check for the stability of the results, calculations with the first model were iterated three times using the method of successive averages to smooth the matrices. Table 3 presents travel times calculated independently for each infrastructure project.

Table 3 is divided in two parts, whereas the upper part lists the changes in public transportation travel times and the lower part the changes in travel times of private transportation. The second column describes travel time changes if there are no modal split changes. The third and fourth column lists the travel times after modal split calculation. The last column presents the results after three iterations between modal split calculation and assignment with parameter set 1. It means that modal split changes are balanced owing to congestion effects in private transportation. The second mode choice parameter set implies less change, partly because it does not capture all relevant changes in generalized costs. The first mode choice parameter set includes additional parameters and can be characterized as more sensitive in this case. Note that travel time increases can be due to reduced capacities (e.g., Project 4) or increased

Table 2
Infrastructure projects considered

Public transportation	*Investment [Mio sFr]*	*Private transportation*	*Investment [Mio sFr]*
5 Upgraded commuter railway system	300	10 Large bypass (south)	1,200
11 Extension of regional train network	270	12 Large bypass (east)	950
8 Tramway 1 (Köniz Schliern)	190	2 Access road (Morillon)	700
9 Tramway 2 (Bern West)	160	4 Removal of existing bypass	450
6 Accelerated express trains	130	14 Access road (Münsingen)	90
7 Higher frequency on regional trains	130	3 Access road (Zollikofen)	90
		13 Expansion of a major junction	60
		1 Small bypass	10

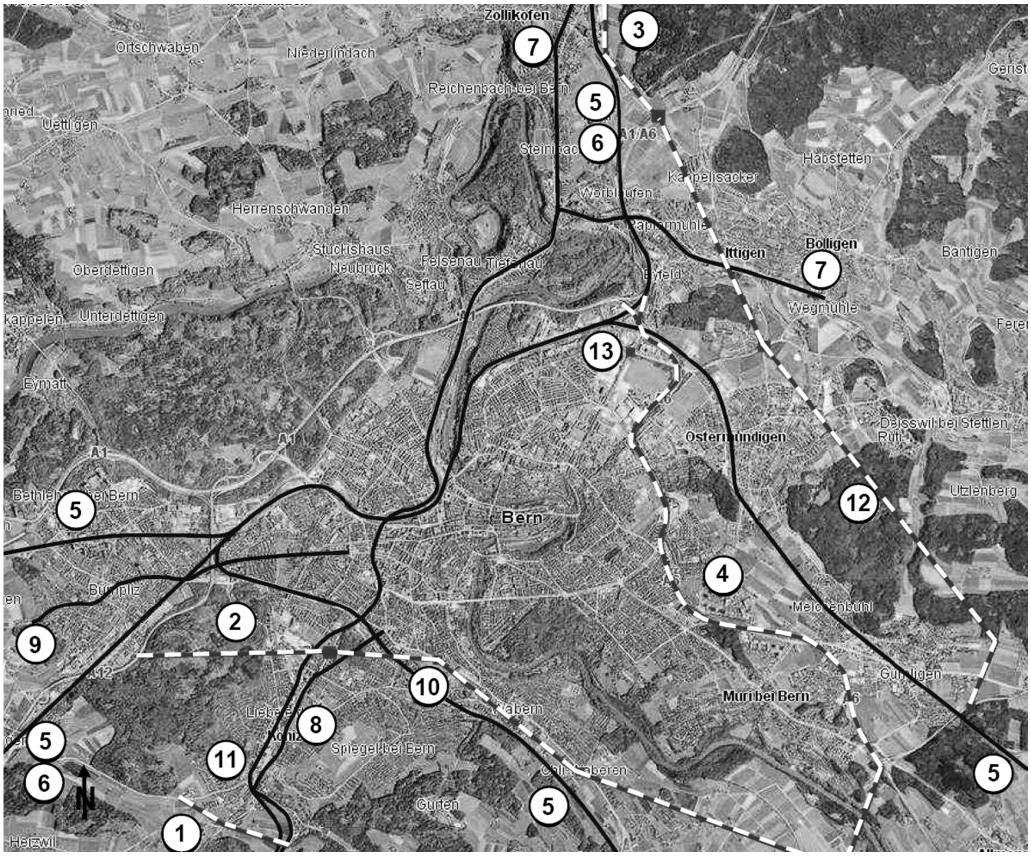

Fig. 2. The region of Bern and the infrastructure projects. Private transportation projects are in *dashed lines; numbers* refer to Table 2.

demand (e.g., Project 5). Changes observable after additional iterations are generally small and do not justify the computational effort during optimization.

4.2 Convergence of the ant colony heuristic

The convergence of the ant colony heuristic is the precondition for its success in application. Setting the parameters α, β, and ρ (P1 and P2) within a neutral range does not guarantee convergence. It depends to a certain extent on the value range of the objective function. Notably, the weights of the variables to calculate probability p_{ij} (P1) cannot be captured with the suggested standard values (see below for further explanations). Very few related papers take notice of this crucial topic (Merkle et al., 2002). We found no consideration of this issue for the NDP in existing literature.

A general problem, also for the calibration issue, is the assessment of the result quality obtained by the ant colony algorithm, as a complete enumeration is infeasible. The relative position vis-à-vis the true optimum bundle cannot be known because of the heuristic nature of the algorithm. However, a manual analysis of iterations indicates the quality of the obtained result. Convergence behavior can be verified to a certain extent using graphics like those explained later (Figures 3 and 4). First, the ant tours are composed of projects more randomly. Then, there has to be a continuous increasing amount of pheromones on the relevant links. As mentioned earlier, there is no possible way to prove that the algorithm has reached the global maximum. When applying methods to avoid local optima, the possibility of reaching the global maximum can be maximized.

According to Poorzahedy and Abulghasemi (2005), convergence behavior improves when the ants start at each project once, before pheromone markers are recalculated. As our specific case study contains 14 different infrastructure projects, the pheromone markers are updated only after these 14 bundles. This seems to be the general rule for the pheromone update: for more complex combinatorial problems, better results are obtained with less frequent pheromone updates (Dorigo and Stützle, 2004). Results also improve when ants start

Table 3
Travel time change due to the projects in isolation with and without modal choice effects

No.	*Project*		*After assignment*	*One iteration: second parameter set*	*One iteration: first parameter set*	*Three iterations: first parameter set*
Change in public transport travel times [h]						
1	Small bypass	Private	0	−180	−310	−299
2	Access road (Morillon)	Private	0	−34	−262	−253
3	Access road (Zollikofen)	Private	0	−146	−295	−291
4	Deconstruction of existing bypass	Private	0	143	−32	−30
5	Upgraded commuter railway system	Public	−311	85	311	289
6	Accelerated express trains	Public	−683	−995	1,637	1,585
7	Higher frequency on regional trains	Public	−206	40	−281	−261
8	Tramway 1 (Köniz Schliern)	Public	−28	10	−229	−227
9	Tramway 2 (Bern West)	Public	−14	8	−254	−248
10	Large bypass (south)	Private	0	−649	−543	−524
11	Extension of regional train network	Public	−209	24	117	110
12	Large bypass (east)	Private	0	−154	−396	−388
13	Extension of main junction	Private	0	−110	−275	−266
14	Access road (Münsingen)	Private	0	−30	−275	−274
Change in private transport travel times [h]						
1	Small bypass	Private	−144	−89	−229	−235
2	Access road (Morillon)	Private	−33	−32	−161	−145
3	Access road (Zollikofen)	Private	−98	−89	−193	−199
4	Deconstruction of existing bypass	Private	504	423	339	267
5	Upgraded commuter railway system	Public	0	−125	−339	−349
6	Accelerated express trains	Public	0	−86	−650	−648
7	Higher frequency on regional trains	Public	0	−79	−151	−113
8	Tramway 1 (Köniz Schliern)	Public	0	−12	−155	−127
9	Tramway 2 (Bern West)	Public	0	−19	−147	−111
10	Large bypass (south)	Private	−867	−737	−903	−891
11	Extension of regional train network	Public	0	−87	−252	−250
12	Large bypass (east)	Private	−305	−302	−400	−381
13	Extension of main junction	Private	−93	−51	−182	−181
14	Access road (Münsingen)	Private	−57	−78	−191	−171

in strict rotation, instead of choosing the initial project randomly (Poorzahedy and Abulghasemi, 2005).

The benefits (mean and variance) of bundles tested during a run are similar for various runs. Below, a single run is used for illustration; calculations were executed using the second modal split parameter set and enhanced parameter settings (see the following section for details). Convergence behavior is shown in Figure 3, and statistics of the benefit–cost ratios during the run are shown in Figure 4. Figure 3 shows the pheromone amounts of each link, corresponding to the links in Figure 1. The gray color of the squares is proportional to the pheromone amount. There is no separation of directions on the links; therefore, the squares are located only on one side of the diagonal. Most of the squares diminish while a few stay stable or even get darker. It is possible to recognize relevant links after the seventh iteration. The squares belonging to the selected bundle never appear in black because of the constant evaporation rate. Convergence has to take place continuously. If convergence would have been reached in one of the first iterations, a local maximum can be assumed, making it possible to identify misleading parameter sets and early convergence with graphics like Figure 3. Convergence behavior should be analyzed carefully to verify the results; using a graphic like Figure 3. No convergence would lead to different random patterns.

Figure 4 shows the benefit–cost ratio of the evaluated bundles for a complete run. The average benefit–cost ratio increases at the beginning and remains stable toward the end. The heuristic finds the most favorable bundle first during the third iteration. After that, the best bundle is chosen more and more frequently. At the end of a complete run, the best bundle is reached by nearly each ant tour. The combination of links, which belongs to

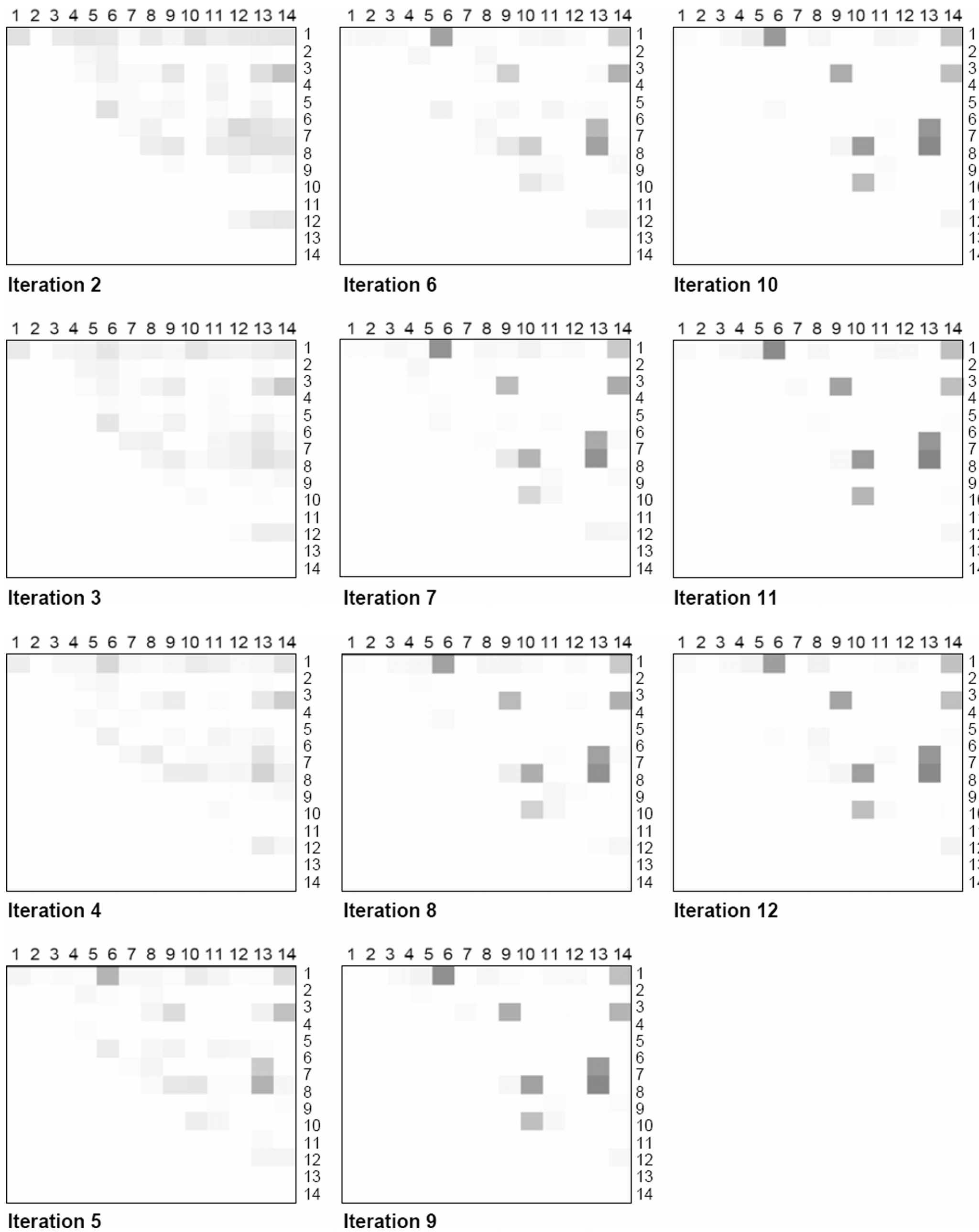

Fig. 3. Pheromone marker density during a complete optimization run. (The numbers correspond to the projects.)

the best bundle, has a very high pheromone density and therefore, the corresponding probability of choosing the same combination is very high as well. However, there are always bundles with lower benefit–cost ratios, even when the algorithm has already converged. They occur particularly when ant has to start at projects with low benefit–cost ratios, because the ant starts at each project once (see Section 2.3). At first, the random choice of the next project is the reason for the unfavorable solutions. Later, the algorithm always finds the most favorable bundle as long as the ant starts at a project belonging to the best bundle. In Figure 4, oscillation of mean benefit–cost ratio can be seen until the end of the run. This is due to remaining variance of the bundles tested in each iteration.

Using Figures 3, 4, and the list of bundles selected and benefits calculated, one can identify when convergence has been reached. Distinct squares (in Figure 3) have to match the convergence approach (Figure 4) and the largest benefits.

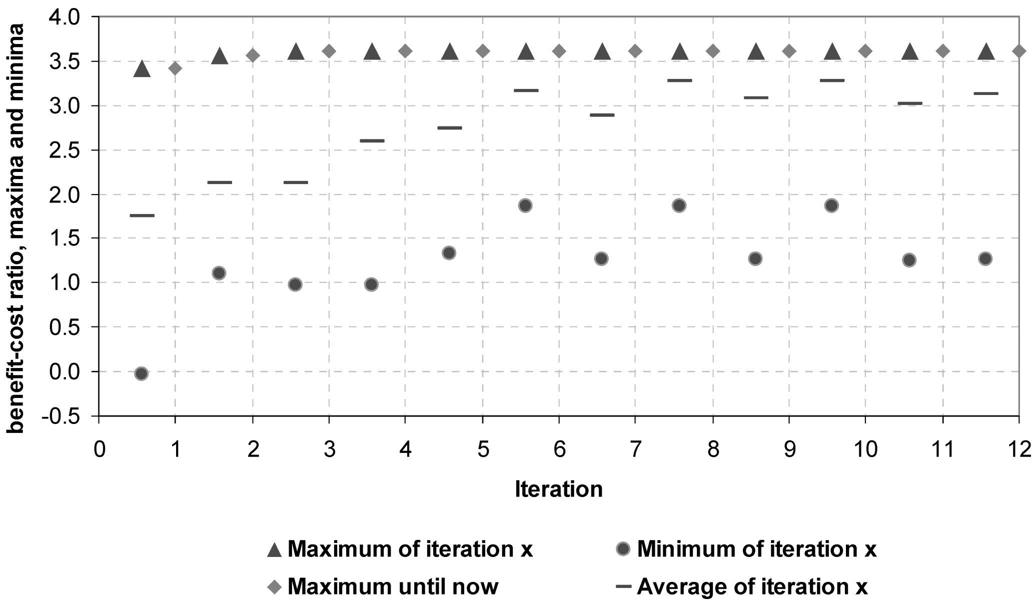

Fig. 4. Benefit–cost ratios during a complete run.

4.3 Calibration of ant algorithm parameters

Calibration has to be carefully performed, employing a systematic approach to reduce the number of equilibria calculated and computation time. The following methodology is straightforward and considers only the three most important parameters (Table 4). They all manipulate the influence of pheromone markers during the ants' decision-making process. At least one complete run has to be performed for each parameter set, such as that shown in Figure 4 above (one run with 12 iterations and 14 times 12 project bundles, or benefit–cost calculations, respectively). At the beginning, we recommend systematic parameter settings to find the most efficient set as quickly as possible (Figure 5). It is desirable to have as few runs as possible to minimize computing times. It is also recommendable to analyze results during calibration so false parameter sets can be detected whenever no convergence occurs during a complete run.

The parameter sets, which were tested here, are displayed in Figure 5. A neutral position is adopted in the center with $^{\alpha}/_{\beta} = 1$ and $\rho = 0.5$, as proposed by Poorzahedy and Abulghasemi (2005). Neutral means that the relevant parameters have the same influence within the decision process of the ants. So, the pheromone concentration τ_{ij} and the benefit of the infrastructure project N_j have the same weights when calculating the probability p_{ij}. In addition, during recalculation of the pheromone concentration, the average is calculated of the already existing pheromones and the new pheromones of the latest bundle ($\rho = 0.5$). The initial four parameter sets are arranged in a rectangle around

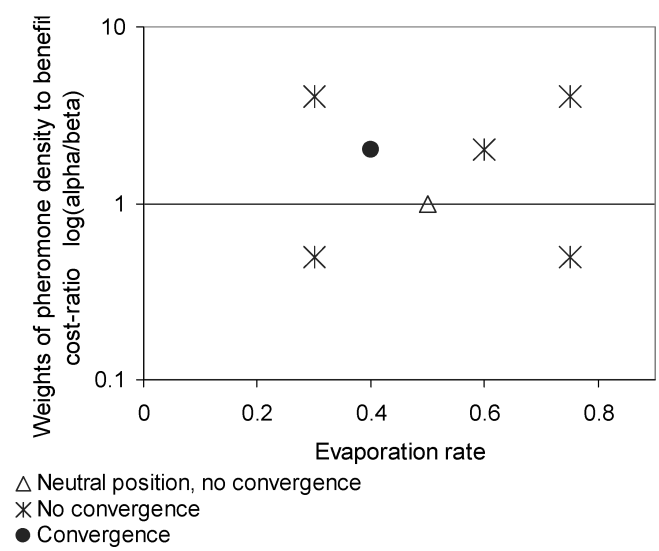

Fig. 5. Different parameter sets tested.

Table 4
Three parameters to be calibrated

Parameter	*Explanation of parameter*	*Corresponding formula*
α	Weight of pheromone concentration	(P1)
β	Weight of an individual infrastructure project	(P1)
ρ	Evaporation rate	(P2)

the neutral position. Thus, a large area can systematically be covered and the chance increases for successful and fast convergence behavior. The corners of the rectangle efficiently show the effect of the parameters. Please notice that each parameter setting requires a complete run, like that shown in Figure 4, which means an equivalent amount of traffic equilibrium assignment.

When no set delivered convergence, the search was extended using combinations inside the box defined by the first four sets.

The lower and upper bound of $^{\alpha}/_{\beta}$ were chosen—on one hand—to limit the influence of the pheromones (small α), so that no "learning effect" will occur. On the other hand, when α is too large, the first bundle with a good result will be chosen by the following ants and convergence will occur too quickly in a local optimum. Calibration of the evaporation rate ρ is less sensitive, but important as well. A higher evaporation rate focuses more on recent results and shortens the feedback loop. We started at $\rho = 0.5$ and improved the results when we applied higher evaporation rates. The ability to lose information gained after a certain period of time is equivalent to "forgetting" in artificial intelligence research.

In addition, properties of the exponential function have to be considered for enhanced convergence. The smaller the value of α and β, the larger the random effect of the probabilistic choice function. Results of the calibration run, especially the values of each bundle, have to be considered very carefully to identify convergence behavior. In this case, the ratio $^{\alpha}/_{\beta}$ cannot be lower than 1 because no convergence was recognized. However, calculations with a ratio of 4 ended up in a local optimum, recognized by analyzing bundles and the corresponding benefit–cost ratios.

The selected parameter set could be improved with additional calibration runs, but there is a trade-off between additional calibration effort and final results obtained. Nonetheless, the discrete nature of the problem and the relatively small number of projects make an improvement unlikely. In addition to the calibration above, it is possible to change parameter values during an optimization run to obtain better convergence performance (see later). It is essential to recalibrate the parameters whenever changes are made in the modal split or objective function. Here, such a recalibration was performed after an earlier change in the modal split function.

4.4 Comparison between the projects and the evaluated bundle

The two modal choice parameter sets have a surprisingly large impact on the chosen bundle. The private car-oriented projects do not vary between the two optimal bundles, but the transit projects are almost entirely different (Table 5). Table 6 shows the benefit–cost ratios of all infrastructure projects, calculated individually. The small bypass (Köniz) has the highest benefit–cost ratio due to very low building costs. The reconstruction (downgrading of its function and capacity) of the existing bypass reduces capacity of the eastern highway and results in a negative cost–benefit ratio because environmental variables are not included in the assessment here. So, its emission reduction objectives are not fully valued in this densely populated area.

Generally, public transportation projects show lower ratios. There are several possible reasons for this outcome. First, projects could, in fact, be less efficient than private transportation projects. Second, the benefit–cost

Table 5
The selected bundles by mode choice parameter set employed

Mode	*First parameter set*	*Second parameter set*
Public transportation		
	Tramway 1 (Köniz Schliern)	Accelerated express trains
	Tramway 2 (Bern West)	Tramway 2 (Bern West)
		Higher frequency of regional trains
Road projects		
	Access road (Münsingen)	Access road (Münsingen)
	Access road (Zollikofen)	Access road (Zollikofen)
	Expansion of main junction	Expansion of main junction
	Large bypass (south)	Large bypass (south)
	Small bypass (Köniz)	Small bypass (Köniz)

Please take notice of the assessment and the small number of indicators calculated for the benefit–cost ratio. Current evaluations of projects within the region of Bern could lead to different results.

Table 6
Benefit–cost ratios of the projects in isolation (first modal split parameter set)

Projects	*Mode*	*Costs/year (Mio. CHF/a.)*	*Travel time savings/year (Mio. CHF/a.)*	*Benefit–cost ratio (First year return)*
Small bypass	Private	0.4	68	169
Extension of main junction	Private	2.3	57	25.2
Access road (Zollikofen)	Private	3.5	68	19.4
Access road (Münsingen)	Private	3.5	59	16.8
Higher frequency on regional train	Public	5.0	44	8.8
Tramway 2 (Bern West)	Public	6.3	47	7.5
Tramway 1 (Köniz Schliern)	Public	7.5	50	6.7
Extension of regional train network	Public	10.8	51	4.7
Upgraded commuter railway system	Public	13.9	52	3.7
Large bypass (south)	Private	45.0	175	3.7
Large bypass (east)	Private	37.5	98	2.6
Access road (Morillon)	Private	27.5	48	1.7
Accelerated express trains	Public	5.0	7	1.4
Removal of existing bypass	Private	17.3	−46	−4.6

Please take notice of the assessment and the small number of indicators (e.g., missing environmental indicators); additional indicators could lead to different results.

function could be incomplete, due to considering only travel times and operation costs and lack of other impacts, such as environmental and social benefits. Neither modal split parameter set covers all aspects of mode choice fully, for example, comfort of vehicles or reliability. Substituting a tramway for a bus line (like the two tramway projects considered) can result in low travel time savings and therefore a low benefit–cost ratio, even though total benefit may be higher.

4.5 Network consequences

All included projects operate at full capacity. Applying the ant colony heuristic, possible interactions between projects are taken into consideration. Traffic decreases are remarkable on notoriously congested links such as motorway sections west and east of the city. Referring to detailed network analysis, traffic decreases are due to the new bypass, and the expansion of the major junction, respectively. Traffic also decreases on permanently congested streets in the center of the city.

Table 7 shows total travel times when comparing the evaluated bundles with the sum of the isolated projects, separately for the two selected bundles. One sees that cumulated travel times of single projects are substantially greater than travel times of bundles, because the projects compete for the same users. The method employed ensures that these interactions are identified and properly accounted for.

Table 7
Comparison of the projects in isolation with the bundles (both evaluated with the first modal split parameter set)

	Reference scenario total travel time	*Optimal bundle savings*	*Sum of isolated projects savings*
Bundle of first modal split parameter set			
Public transport [Passenger.h]	58,483	2,000	2,129
Private transport [Veh.h]	46,151	999	2,000
Value [Mio. sFr/a]		226	524
Bundle of second modal split parameter set			
Public transport [Passenger.h]	58,483	555	844
Private transport [Veh.h]	46,151	1,153	1,228
Value [Mio. sFr/a]		218	351

5 CONCLUSIONS

The article has demonstrated that a substantial NDP for a large network can be solved with a reasonable effort employing the ant colony heuristic. The best bundle of projects cannot be compiled only considering the isolated benefit–cost ratio of the single projects because of the interactions between infrastructure projects. The algorithm successfully copes with this problem and takes account of possible interaction between projects. The results also show the sensitivity of the final results to the overall modelling framework, here exemplified by the different modal split approaches. It is clear that the incorporation of both destination choice and trip generation, as for example in the Swiss National model (Vrtic et al., 2007), might change the bundles again.

The computation times implied in such a complete equilibration require a faster approach. It will also be necessary to remove or reduce the need for the parameter calibration, or at least to automate it. It is highly recommended to use a similar systematic approach for parameter calibration when applying the algorithm to a large network. The integration of the benefit–cost calculation (naturally a richer one than implemented here), into the overall software environment would be highly desirable. In addition, the methods to avoid local optima should be improved to reduce computation time and manual checking (Russo and Vitetta, 2006).

Fixed values of α, β, and ρ could be adjusted during the run to improve both speed and quality of the solution. The parameters could be increased or scaled down during the course of a run. For example, at the beginning, α could be smaller, so that the ants choose projects more at random; evaporation would also be small. After a few iterations, α would grow, so that the influence of the pheromone markers would increase during the decision-making process. The ants would start to benefit more intensively from the preceding experiences.

Another important issue is the fixed budget restriction. It is possible to achieve higher cost–benefit ratios when, for example, not using the entire budget. The algorithm should be adapted so that it is possible to stop before the budget is fully committed.

Finally, the staging of the projects should be addressed in conjunction with an appropriate modelling of the changes in travel demand and population and work place distribution. It is clear that the complexity of this programme is still beyond current computing capabilities, but research into this issue should start immediately.

ACKNOWLEDGMENTS

The authors thank the Canton of Bern and Mr. Ulrich Seewer for making the network model available. Rapp Trans AG, Basel, and in particular Mrs. Graciela Christen and Mr. Simon Kettner provided support and advice with access to EMME and their knowledge of its scripting language. The basic idea for this article was developed jointly with Mr. Rudolf Burger of R+R AG, Baden. This article expands on Vitins (2006) written under the supervision of the coauthor.

Any errors are our own.

REFERENCES

Axhausen, K. W., Hess, S., König, A., Abay, G., Bates, J. J. & Bierlaire, M. (2008), Income and distance elasticities of values of travel time savings: New Swiss results, *Transport Policy*, **15**(3), 173–85.

Bakuli, D. L. & Smith, J. M. C. (1996), Resource allocation in state-dependent emergency evacuation, *European Journal of Operational Research*, **89**, 543–55.

Bell, M. G. H. (2000), A game theory approach to measuring the performance reliability of transport networks, *Transportation Research Part B*, **34**(3), 533–45.

Bonabeau, E., Dorigo, M. & Theraulaz, G. (2000), Inspiration for optimization from social insect behaviour, *Nature*, **406**(6), 39–42.

Braess, D. (1968), Über ein Paradoxon aus der Verkehrsplanung, *Unternehmensforschung*, **12**, 258–68.

Dorigo, M., DiCaro, G. & Gambardella, L. M. (1999), Ant algorithms for discrete optimization, *Artificial Life*, **5**(2), 137–72.

Dorigo, M., Maniezzo, V. & Colorni, A. (1996), The ant system: optimization by a colony of cooperating agents, *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, **26**(1), 1–13.

Dorigo, M. & Stützle, T. (2004), *Ant Colony Optimization*, MIT Press, Cambridge.

Frank, M. & Wolfe, P. (1956), An algorithm for quadratic programming, *Naval Research Logistics Quarterly*, **3**, 95–100.

Garay, M. R. & Johnson, D. S. (1979), *Computers and Intractability—A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, San Francisco.

Hsieh, T.-Y. & Liu, H.-L. (2004), Genetic algorithm for optimization of infrastructure investment under time-resource constraints, *Computer-Aided Civil and Infrastructure Engineering*, **19**(3), 203–12.

INRO (1998), *EMME/2 User's Manual, Software Release 9*, INRO Consultants Inc., Montréal.

Keeney, R. L. & Raiffa, H. (1976), *Decisions with Multiple Objectives*, Wiley, New York.

Kellerer, H., Pferschy, U. & Pisinger, D. (2004), *Knapsack Problems*, Springer-Verlag, Berlin.

LeBlanc, L. (1975), An algorithm for the discrete network design problem, *Transportation Science*, **9**(3), 183–99.

Merkle, D., Middendorf, M. & Schmeck, H. (2002), Ant colony optimization for resource-constrained project

scheduling, *IEEE Transactions on Evolutionary Computation*, **6**(4), 333–46.
Optima Consortium. (1997), *Report on Work Packages 30 and 40*, Institute for Transport Studies, University of Leeds, Leeds.
Poorzahedy, H. & Abulghasemi, F. (2005), Application of ant system to network design problem, *Transportation*, **32**(3), 251–73.
Russo, F. & Vitetta, A. (2006), A topological method to choose optimal solutions after solving the multi-criteria urban road network design problem, *Transportation*, **33**(4), 347–70.
RVK4 & Rapp Trans. (2002), *GVM Region Bern: Etablierung von Modal Split-Funktion*. Basel, Rapp Trans.
RVK4, VRB & Rapp Trans. (2004), *Gesamtverkehrsmodell der Region Bern*. Basel, Rapp Trans.
Spiess, H. (1984), Contributions à la théorie et aux outils de planification des résaux de transport urbain, *Ph.D. thesis*. Department of Computer Science and Operational Research, University of Montreal.
Spiess, H. & Florian, M. (1989), Optimal strategies: a new assignment model for transit networks, *Transportation Research B*, **23B**(2), 83–102.
Stopher, P. R. & Meyburg, A. H. (1976), *Transportation System Evaluation*, Lexington Books, Lexington.
Sumalee, A. (2007), Multi-concentric optimal charging cordon design, *Transportmetrica*, **3**(1), 41–71.
Sumalee, A., Watling, D. P. & Nakayama, S. (2006), Reliable network design problem—case with uncertain demand and total travel time reliability, *Transportation Research Record*, **1964**, 81–90.
Taber, J. T., Balling, R., Brown, M. R., Day, K. & Meyer, G. A. (1999), Optimizing transport infrastructure planning with a multiobjective genetic algorithm model, *Transportation Research Record*, **1685,** 51–56.
Tuydes, H. & Ziliaskopoulos, A. (2006), Tabu-based heuristic approach for optimization of network evacuation contraflow, *Transportation Research Record*, **1964,** 157–68.
Vitins, B. J. (2006), Optimierung der Verkehrsinfrastruktur der Region Bern mit der Ant Colony Heuristic, *Diplomarbeit*, IVT, ETH Zürich, Zürich.
Vrtic, M., Fröhlich, P., Schüssler, N., Axhausen, K. W., Schulze, P., Kern, P., Perret, F., Pfisterer, S., Schultze, C., Zimmermann, A. & Heidl, U. (2005), *Verkehrsmodell für den öffentlichen Verkehr des Kanton Zürich*, Bericht für das Amt für Verkehr, Kanton Zürich, IVT, Ernst Basler + Partner, Zürich and PTV, Karlsruhe.
Vrtic, M., Lohse, D., Fröhlich, P., Schiller, C., Schüssler, N., Teichert, H. & Axhausen, K. W. (2007), A simultaneous two-dimensionally constraint disaggregate trip generation, distribution and mode choice model: theory and application for a Swiss national model, *Transportation Research A*, **41**(9), 857–73.
VSS. (2006), *SN 641 820 Kosten-Nutzen-Analysen (KNA) bei Massnahmen im Strassenverkehr*, VSS, Zürich.
VSS. (2007), *SN 641 821 Diskontsatz in Kosten-Nutzen-Analysen im Verkehr*, VSS, Zürich.
Wardrop, J. G. (1952), Some theoretical aspects of road traffic research, *Proceedings of the Institute of Civil Engineers*, **1**(Part II), 325–78.