Transportation Research Interdisciplinary Perspectives 27 (2024) 101222

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Sensitivity analysis of activity scheduling parameters with a parameter optimization framework

Matheus Moro Zamprogno, Domokos Esztergár-Kiss *

Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, 1111 Budapest, Műegyetem rkp 3, Hungary

A R T I C L E I N F O

Keywords:
Genetic algorithm
Sensitivity analysis
Evaluation framework
Parameter setting

A B S T R A C T

Transportation-related activity scheduling is becoming more complex due to the growing number of potential locations and extensive opportunities to visit various places. Throughout the years, in the field of transportation several attempts were made to optimize travelers' activity chains with different parameters to set, but there is a lack of comprehensive solutions. In this research, the activity chain optimization algorithm is applied, which requires high computational efforts. To provide an adequate calibration of the parameters, a sensitivity analysis is conducted. The aim of the analysis is to reveal how changes in the attribute values modify the final outcomes. The relevant parameters, activity chains, transport modes, optimization algorithms, and fitness functions, are identified and considered. For each parameter, an investigation is conducted to reveal its behavior throughout the runs. For example, changes in the population size and crossover function lead to more reliable results, while alteration in the number of generations and the mutation function have no effects on the outcomes. The analysis presents a peculiar behavior of the parameters related to the activity chains. The results can be useful for transportation planners and service providers in the adaptation of the existing network and transportation services to the travelers' mobility patterns.

1. Introduction

Problems in urban transportation networks are frequently related to the accelerated urbanization and growth of cities. As urban locations are becoming more and more attractive to people, the overload of transportation networks becomes an urgent issue (Sinha, 2003). Therefore, network capacity has a direct impact on urban development and the quality of life. Thus, efficient ways of planning are required to maintain a sustainable usage of the existing transportation network (Camargo Pérez et al., 2015).

A smart way of using the transportation network efficiently is by exploiting its capacity and optimizing the actual trips conducted within the system. Such exploitation is well-represented by finding the best routes or closest locations when visiting a city while considering user specific constraints and other options. Throughout the years, several authors addressed the routing problem in different cases while proposing diverse solutions to the issue. Suggestions for optimal bus transit routes and multimodal trip routing are examples of solutions in the field of public transport (Fan and Machemehl, 2006); (Bast, 2016) . Moreover, there are investigations in freight transportation including time-window constraints and multimodal solutions, as well (Yang et al., 2015); (Fazayeli et al., 2018) .

Solving routing problems involves a considerable computational cost. Traveling from one location to another in urban environment requires an analysis of several routes and constraints, which may result in a huge number of calculations. Therefore, advanced research solutions analyze alternative algorithms that could reduce the computational costs of solving routing problems. One of these heuristic algorithms is the genetic algorithm (GA) family and its variations. For example, Beed et al. (Beed et al., 2017) apply the GA to solve the traveling salesman problem (TSP), while Esztergár-Kiss et al. (Esztergár-Kiss et al., 2018) use the GA to optimize travelers' activity chains. Although several solutions appear, the efficiency of the solutions and the setting of suitable parameter need more analysis.

To find solutions for the routing problems and to optimize the usage of the transportation network capacity, the concept of activity chain optimization (ACO) can be used, which describes the tasks done and the locations to visit during a period in urban environment. These tasks can be created artificially, as demonstrated by Charypar and Nagel (Charypar and Nagel, 2005), or they can rely on actual activities, as

* Corresponding author.
E-mail address: esztergar@mail.bme.hu (D. Esztergár-Kiss).

https://doi.org/10.1016/j.trip.2024.101222
Received 8 November 2023; Received in revised form 30 May 2024; Accepted 10 September 2024
Available online 14 September 2024

addressed by Esztergár-Kiss et al. (Esztergár-Kiss et al., 2018). The latter framework represents people's routines more adequately, and it can provide an optimization process that presents routing solutions by involving the optimal set of activities.

The optimization process requires the understanding of all parameters included in the problem. It is fundamental to know how the travel time and transport mode influence the results of the optimization. In the ACO problem, several parameters are present; thus, the problem arises how all these parameters can be set to provide the most suitable solution. Therefore, the main research question is the following: how do the parameters impact the output of the optimization process? For an adequate calibration of the method, sensitivity analysis could be applied. Sensitivity analysis is a valuable method to investigate various parameters.

For this reason, current research work describes the sensitivity analysis of the parameters in the ACO framework. The aim of the study is to perform a detailed parameter analysis and present the most suitable settings for a specific transportation problem. Additionally, this research investigates the impacts of the main attributes on the outcomes of the optimization. The paper is structured as follows. After the introduction in Section 1, relevant literature about sensitivity analysis methods is discussed in Section 2. Section 3 demonstrates the description of the problem and the method applied during the sensitivity analysis. Afterward, in Section 4, the results are presented. The outcomes are discussed in Section 5, and finally, the conclusion is demonstrated in Section 6.

2. Literature review

Relevant papers that apply specialized algorithms to solve routing and optimization problems are introduced with a specific focus on sensitivity analysis.

2.1. Routing problems

Bast et al. (Bast, 2016) present a review of practical algorithms used in route planning including schedule-based public transport and multimodal options. The authors state that routing with public transport is a significantly hard problem due to timetables and multicriteria issues. In the paper, shortest path algorithms on static networks and their performance in real road networks are compared by using query time and processing time. However, specific parameter settings and sensitivity analysis are not realized.

An example of using the GA to solve routing problems is found in the study of Fan and Machemehl (Fan and Machemehl, 2006). The researchers suggest an approach for examining an optimal bus transit route with variable demand. Furthermore, a sensitivity analysis of the main parameters is conducted, which makes the refinement of the algorithm possible thus giving more adequate routes. However, in this case a relatively simple routing problem is solved with a limited number of parameter settings. Regarding the vehicle routing problem, Yang et al. (Yang et al., 2015) propose a solution with time windows by using a hybrid GA. The researchers provide a solution to the heterogeneous vehicle fleet problem by considering the carbon effects. Suggestions for routes and load capability are provided by conducting a sensitivity analysis on the objective parameters. The study focuses on a logistics application where different parameters are relevant than in case of the ACO since in the ACO framework, passenger transportation-related issues are handled. Similarly, Fazayeli et al. (Fazayeli et al., 2018) use the GA to introduce a location-routing solution in a multimodal transportation network. The algorithm is meant to decide on the used transport modes and node locations as well as on the routing and the depot placement. A distribution system is considered where time window constraints are introduced along with fuzzy variables for the representation of the customer demand. Similarly to the previous paper, the aim is to distribute the products of a supply chain, which faces different challenges than organizing travelers' activities in a city.

2.2. Optimization problems

The activity chains of transportation users can be described as a series of ordered activities realized by a person during a day (Yin et al., 2021). By considering the optimization of these activity chains, Charypar and Nagel (Charypar and Nagel, 2005) apply the GA to generate an all-day schedule, which is used to create a multi-agent model for transportation planning. The GA initiates a population of possible activities and selects the best fit for human behavior through the genetic operators. The applied model generates daily activity schedules, but the performance of the algorithm and the settings of the parameters are not evaluated within the analysis. Moreover, Beed et al. (Beed et al., 2017) model and solve a multi-objective TSP problem by using weighted sums and the GA. Sensitivity analysis is conducted to investigate the impacts of various genetic operator values with different weights on the results of the fitness function. However, the algorithm does not reflect the complexity of the ACO problem, and several related parameters are neglected. To solve the TSP problem, Esztergár-Kiss et al. (Esztergár-Kiss et al., 2018) propose an optimization algorithm applied to a real transportation network with timetable data while using temporal and spatial flexibility. The method considers the flexibility of activities and processing times requiring the introduction of the GA with a point of interest (POI) search algorithm. However, no sensitivity analysis, which would support the parameter setting of the problem, are conducted. Moreover, Esztergár-Kiss et al. (Esztergár-Kiss et al., 2020) provide a solution including multimodal features and time window constraints. Additionally, Esztergár-Kiss (Esztergár-Kiss, 2020) presents how to group travel-related parameters of various activity chains into classification parameters and optimization parameters. The research explores a wide range of parameters and demonstrates useful results for defining the parameters included in current study related to the sensitivity analysis.

The effectiveness of GAs is investigated by several authors. Hassanat et al. (Hassanat et al., 2019) review the literature on the topic of choosing the right values for genetic operators and apply a new dynamic method for setting the parameters. However, parameter tuning is only used for the general TSP problem not covering all necessary parameters. Ulukok (Ulukok, 2017) applies the GA in the domain of numerical optimization. The researcher tests the performance of the algorithm with various parameter settings. The outcomes show that getting an optimum solution with GA requires distinct parameter settings.

2.3. Sensitivity analysis

Previous research shows that the performance of the algorithm strongly depends on the parameter settings. Thus, running a sensitivity analysis is required to provide optimal solutions. The impact of the algorithm parameters on the model can be assessed by sensitivity analysis, which aims to quantify the relative importance of the input parameters in determining the value of the output variables.

The sensitivity analysis explores how varying one input influences the output, which is called the one-factor-at-a-time (OFAT) method, as shown by ten Broeke et al. (ten Broeke et al., 2016). In the literature, there are several methods conducting sensitivity analysis, where depending on the type of the base model, different approaches are suggested. As examples of diverse applications, Fan and Machemehl (Fan and Machemehl, 2006) perform the OFAT method with a base set of parameters to investigate the sensitivity of the genetic operators and the parameters of a specific function on the outputs. Moreover, Dass and Namin (Dass and Namin, 2020) apply the sensitivity analysis to investigate evolutionary algorithms. The researchers vary one parameter at a time to uncover the influence of parameter tuning in the GA. Although, current research applies sensitivity analysis, it does not solve the ACO problem and does not assess the suitable parameters for a scheduling problem in the field of transportation.

This study is conducted by using the OFAT method (ten Broeke et al., 2016), which enables the calibration of the ACO parameters based on the results of the sensitivity analysis. The choice of the method is based on the interaction of the inputs that is not a primary issue, while the focus of the research is a robust observation of changes. The method can reveal the qualitative characteristics of the system, it is computationally cheap and can be used with relatively low effort.

3. Method

Since activity scheduling helps users to optimize their trips in a city, in general, the target group of the research includes users. To solve the ACO problem including the traveling salesman problem with time windows (TSP-TW), a system is developed in Python language. For the development, the GA is used combined with Java environment by running the routing engine of the open trip planner (OTP). To analyze the influence of the parameters on the ACO results, a sensitivity analysis, which can be served as the basis for further parameter calibration, is conducted.

3.1. Aco

The ACO problem can be stated as a directed graph where each edge has a cost for traveling from one location vertex to another. Time windows are assigned to the vertexes, and the objective is to find the minimum cost Hamiltonian cycle with the arrival time within or before the time window (Esztergár-Kiss and Remeli, 2019).

The flexible TSP-TW differs from the TSP in two ways. First, by including the time dimension in the calculations, time windows are considered as constraints for the selection of the locations. Second, the TSP-TW includes the option of spatially and/or temporally flexible activities, which can be replaced by other activities of the same type but with a lower total cost (Esztergár-Kiss and Rózsa, 2015).

The TSP-TW is implemented inside the ACO framework, which consists of a series of processes to optimize the activity chain. The first step is the definition of the data input regarding the characteristics of the activity. Afterward, the alternative activity locations are searched with the OTP in case of spatially flexible locations. Additionally, the OTP calculates a travel time matrix for various transport modes. The flexible TSP-TW is solved, and the framework chooses the optimal value. For comprehension reasons, the steps of the flexible TSP-TW method are summarized:

- **Defining the daily activity chain:** A list of activities is defined by the user.
- **Solving the basic TSP-TW:** The original locations are calculated.
- **Prioritizing the activities:** Fixed points receive higher priority than flexible ones.
- **Replacing the flexible points:** New locations are found for the flexible activities within a determined distance.
- **Optimization:** For each version, a TSP-TW is calculated, and the one with the lowest cost is compared with the basic TSP-TW result.

To make the system able to calculate the most adequate schedule in a comprehensive time, a heuristic GA approach is introduced. Otherwise, the optimization would take a long period considering the enormous number of possibilities.

3.2. GA

GAs are heuristic algorithms that can be used for searching, optimizing, and learning tasks. These algorithms were introduced by Holland (Holland, 1975) based on the theory of the survival of the fittest. The crossover and mutation abilities play an important role in continuing the evolutional process. The crossover brings together the "better" traits over time, and the mutation introduces new changes in the mechanism. The algorithm makes a representation of the natural selection regarding the laws of nature, which ensures that the next generation will be "better" than the previous one. Therefore, there are several elements such as population, fitness function, selection, crossover, and mutation described in the following list:

- **Population**: A collection of chromosomes or candidate solutions.
- **Fitness function**: The "better" the fitness, the higher chances of reproduction.
- **Selection**: The choice of the individuals that are to be reproduced.
- **Crossover**: Interchange of genes between chromosomes to create offspring.
- **Mutation**: A random modification in a gene to introduce new attributes.

The main genetic operators of the algorithm are selection, crossover, and mutation. There is a gradual search for the optimal solution evaluated by the fitness value. The optimal solutions are assembled as a group of chromosomes. The individuals who have the desired chromosomes receive "better" scores and pass these chromosomes to the next generations combining them with other desired chromosomes from other individuals. Thus, a population of individuals with more adequate chromosomes is formed (Wirsansky, 2020).

Therefore, the solutions with poor fitness are eliminated. On the other hand, the solutions with high fitness pass their offspring to the next generations until there is no more improvement, and the "best" solution is found. Although the search space is reduced by the algorithm, and a solution can be found way faster, there are some limitations, as well. For example, the need for parameter tuning, intensive computational operations, the chance of premature convergence, and no guaranteed solution are some limitations. Additionally, the algorithm has the risk of getting stuck in a local optimum, as well.

3.3. Sensitivity analysis

A sensitivity analysis is used to assess the behavior of the ACO. The analysis performed in this case is the OFAT described by ten Broeke et al. (ten Broeke et al., 2016). The method consists of determining a base parameter setting and changing one parameter once while maintaining the others fixed. This type of analysis can reveal the influence of a parameter on the output. The goal of the applied sensitivity analysis is to observe how patterns and emergent properties are created. Additionally, other aims of the analysis are to examine the consistency of the emergent properties and to quantify the variability in the outputs. Thus, to deploy the investigation in the ACO system, a working protocol is established based on the steps demonstrated in Fig. 1.

3.3.1. Identifying the relevant parameters

In the work of Esztergár-Kiss (Esztergár-Kiss, 2020), travel-related parameters are presented and grouped into classification parameters and optimization parameters to define a set of inputs, which model the ACO utility function. The ACO system analyzed in this research follows the same structure of the utility function. Therefore, the parameters in common with those presented by Esztergár-Kiss (Esztergár-Kiss, 2020) are picked due to their relevance to the analysis. The classification parameters are related to the specific characteristics of the user, location, or trip. On the other hand, the optimization parameters support the optimization process; thus, they are relevant for the sensitivity analysis.

Besides the weights of the utility function, other parameters are used in the sensitivity analysis. Some of them are related to the algorithm, while others are related to the users' inputs. In this case, the analysis is meant to understand the behavior of the system in different configurations. To present the selected variables, a summary of all relevant chosen parameters with their meanings, default values of the basic setting of the analysis, and brief descriptions are demonstrated in Table 1.

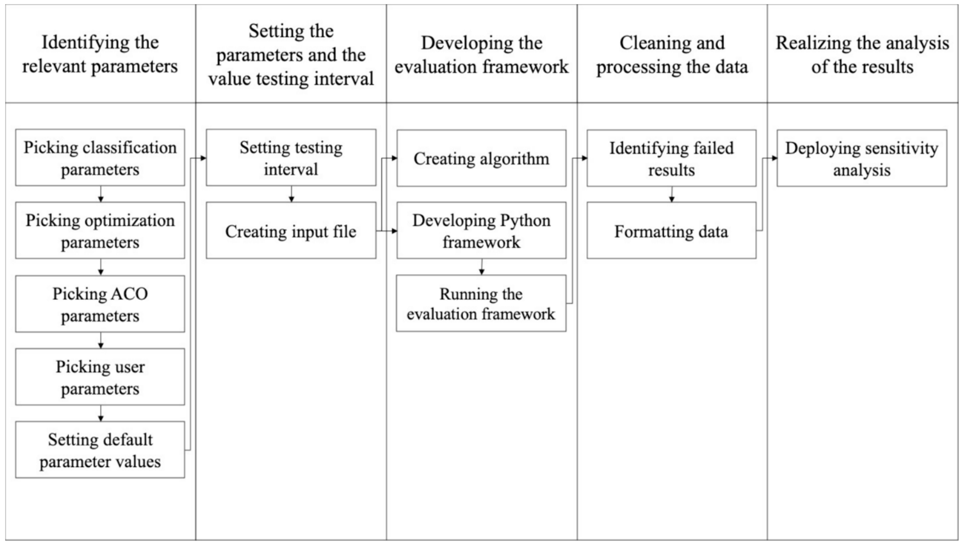

Fig. 1. The flow chart of the applied method.

3.3.2. Setting the parameters and the value testing interval

For defining the testing interval, all parameters are varied uniformly within the possible range as they belong to different terms, and the testing range is heterogeneous among them. Some parameters are tested for the whole search space, e.g., the fitness weights. Other parameters are tested inside a plausible and feasible range, e.g., the start_time. Therefore, it is possible to explore the behavior in a wide space of assumptions, which ensures that interesting behavior is not ignored. Annex 1 demonstrates the testing intervals for all the tested parameters.

It is important to highlight the differences between the values of the problem_solved parameter, which are related to the applied mode of transportation such as car (value: 1), walking (value: 2), public transport (value: 3), and bicycle (value: 4).

The ACO system is meant to read the input file with the user preferences about the activity chain. This file contains the name of the activity, the processing time (spent time), priority label (flexibility), time window open/close, demand time start/end, and the coordinates of the activity. Besides that, all relevant parameters are used in their default value. Moreover, the priority level indicates whether the activity is flexible or not. 1 indicates a fixed activity, 2 stands for a spatially flexible activity, 3 is applied for a temporally flexible activity, and 4 indicates a totally flexible activity. Due to the heuristic behavior of the algorithm, the number of runs for each parameter test is set to 10 to search for the consistency and average performance. In this test, the input file remains constant, as shown in Table 2.

3.3.3. Developing the evaluation framework

Due to the huge number of runs derived from all parameters and their respective testing intervals, as well as the needed simulations for each variable, an evaluation framework capable of running all the tests consecutively is developed. The framework utilizes a second input file where all the parameter values are described to make consecutive runs. The values are read, and one is applied at a time on the correspondent parameter keeping all other parameters in default mode. Afterward, the framework runs the ACO to calculate the results. In the end, the framework saves the results in a file and applies the default values again to all parameters for restarting the process. This process occurs a predetermined number of times (10 times in this case) for each parameter. By the end of all runs, a dataset is built with all the outcomes specifying which result belongs to which parameter modification. Fig. 2 presents the functioning of the framework.

3.3.4. Cleaning and processing the data

After finishing all the runs, the framework returns the compiled results of all simulations in a dataset, but the data are saved as given by the ACO system. Therefore, cleaning unwanted or failed results and processing the ones from which information can be extracted are needed. At the end of the runs, more than 3000 records are collected. After the processing and cleaning of the data, 2220 records remain for the analysis, which is performed by using statistical tools. The main attributes analyzed in the results are the following: travel time, the start and end times of the activities, the order, and the chosen POI.

3.3.5. Realizing the analysis of the results

As mentioned before, the nature of GAs is heuristic. Thus, the consistency of the results is prioritized during the analysis. The prioritization includes the comparison of each attribute to the output by using statistical tools. The range shows the difference between the largest and the smallest values, which demonstrates an initial measure of variability. The sum of the variables is divided by the sample size. In other words, the sample mean is provided, which has the property of describing the balance point of a distribution. Finally, the variance is calculated.

4. Sensitivity analysis results

4.1. Start_time

The outcomes for this parameter shown in Table 3 have similar standard deviations in general, but slightly "better" after the time 480. However, there is an exception regarding the 390-input value, which is with a 14,60 min range, and has the highest deviation.

4.2. End_time

When checking the data, the algorithm chooses six times (out of 10) the same route for the value 1350, which indicates the fastest route of the End_time analysis. Thus, this route is a candidate for being a global optimum. As it can be observed in Table 4, there is an exception regarding the value 1350, which presents a higher deviation compared to the others. Most probably due to a higher local optimum, the second higher result found in the End_time analysis, the deviation becomes higher.

Table 1
Relevant parameters.

Parameter	Meaning	Default value	Description
start_time	The start time of the activity chain	480	User preference In seconds, counting from 00 h = 0 to 23 h59 = 1439
end_time	The end time of the activity chain	1170	User preference In seconds, counting from 00 h = 0 to 23 h59 = 1439
problem_solved	The mode of transportation	0	User preference Each number represents a different mode
population_size	The list of GA individuals	30	Genetic operator parameter The number of candidate solutions
generations	The number of GA generations	20	Genetic operator parameter The number of generations of populations
cxpb	The probability of mating the GA	0.1	Genetic operator parameter The probability of crossing-over solutions
mutpb	The probability of GA mutation	0.2	Genetic operator parameter The probability of changing an element of a solution
number_of_alt_locations	The number of alternative locations kept by the algorithm	5	ACO parameter Alternative places for flexible activities
fitnessweight1	Daily income / cost	0	The weight of the utility function Maximization objective
fitnessweight2	CO2 emission	0	The weight of the utility function Minimization objective
fitnessweight3	Burned calories	0	The weight of the utility function Maximization objective
fitnessweight4	Time	4	The weight of the utility function Minimization objective
fitnessweight5	Subtour by electric vehicle	0	The weight of the utility function Minimization objective
fitnessweight6	Safety	4	The weight of the utility function Maximization objective
fitnessweight7	Comfort	4	The weight of the utility function Maximization objective
fitnessweight8	Quality	4	The weight of the utility function Maximization objective
fitnessweight9	Ownership	4	The weight of the utility function Maximization objective
fitnessweight10	Hot mode	7	The weight of the utility function Maximization objective
fitnessweight11	Cold mode	0	The weight of the utility function Maximization objective
fitnessweight12	Rain mode	0	The weight of the utility function Maximization objective
fitnessweight13	Snow mode	0	The weight of the utility function Maximization objective
fitnessweight14	Humid mode	7	The weight of the utility function Maximization objective
fitnessweight15	Windy mode	0	The weight of the utility function Maximization objective
fitnessweight16	Total time / in-vehicle time	4	The weight of the utility function Minimization objective

4.3. Problem_solved

This parameter corresponds to a different mode of transportation for each parameter value. Thus, it needs specific analysis for each category and for how its change can affect the functioning of the application. Table 5 presents the respective results.

The car (mode 1) represents the default parameter values where it is expected that the trips are made with the car mode. By using this mode, a consistent result appears with a low deviation, as shown in Table 5.

Walking (mode 2) is the slowest transport mode realized by the algorithm. The deviation and range of this mode are rather big compared to the other modes. Due to the low speed, any change in the route can have a huge impact on the output. Table 5 shows a huge difference between the lowest and the highest travel time values.

A combination of public transport and walking (mode 3) is realized because the stops can be solely reached on foot. Thus, the results are quite poor compared to the other modes. According to Table 5, the outputs for these parameters are quite inconsistent with a huge range between the extremes. A possible cause can be that some extreme alternative activity locations are quite far away from the rest of the activities.

The results of bicycle (mode 4) are found between the walking and the public transport modes. Since the speed of a bicycle is higher than the walking speed, the impacts of changing the routes can be smaller, which brings very consistent results (Table 5).

4.4. Population_size

When checking the results, it is possible to identify a tendency in the population outcomes, as presented in Table 6. Thus, the larger the population, the higher the chance for the algorithm to reach “better” results. This result is quite plausible. Once the algorithm has more candidates for possible adequate solutions with a higher population number, the consistency of the outputs is increasing.

Table 2
Default input values.

Activity	Processing time	Priority label	TW open	TW close	Demand time start	Demand time close	Latitude	Longitude
Start	0	1	1	1439	1	1439	47,5405	19,1494
Work	480	1	360	1140	360	1140	47,4806	19,0296
University	120	1	360	1380	420	1380	47,4727	19,0600
Sport	60	3	360	1439	360	720	47,4713	19,0540
Bar & Pub	120	2	600	1439	720	1439	47,4730	19,0576
End	0	1	1	1439	1	1439	47,5405	19,1494

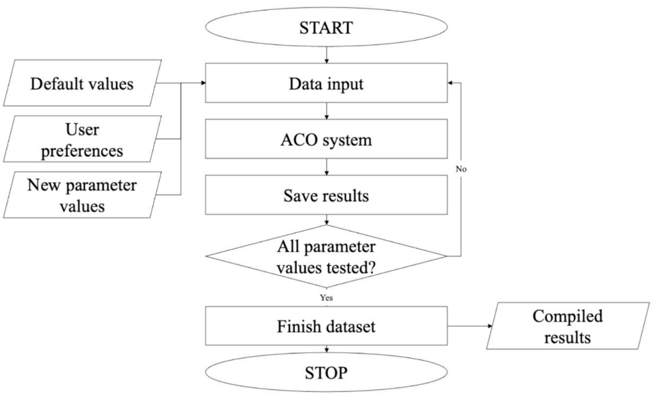

Fig. 2. The functioning process of the framework.

4.5. *Generations*

The outcomes of the tests with the generation parameter, shown in Table 7, demonstrate that there is no tendency regarding the number of generations. The difference between the deviations is rather small, which means that from a low number of generations, it is already possible to obtain adequate results.

4.6. *Cxpb*

The parameter of the crossover probability is quite expressive as it is related to the diversity of the population and the propagation of the "best" results over the generations. In Table 8, it is observable that higher probabilities generate "better" results. It is evident that a 90 % crossover value gives a more adequate result. Furthermore, the values above 50 % have "better" results than the values below 50 %.

4.7. *Mutpb*

Table 9 shows that there is not a clear tendency in the mutation probability parameter behavior over the tests. The highest probability gives the "better" result, but there are adequate results with lower probabilities, as well. Moreover, there could be a possibility that higher mutation rates may disfigure the characteristics of the GA turning it into a random search.

4.8. *Number_of_alt_locations*

This parameter in special can have a direct relationship with the choice of the location as it determines the number of alternative locations to keep for spatially flexible activities. It can be observed in Table 10 that there is a slightly decreasing tendency in the standard deviation with the increase of the number of places. This tendency is obvious as the more options are available for choosing, the higher is the chance of choosing more adequate options.

4.9. *Fitnessweight*

The fitness weights are related to the optimization parameters of the utility function. In this case, an overall analysis is conducted, and it is seen how the parameters perform. In Annex 2, the runs which do not return extreme values and perform "better" among all options with low standard deviation are analyzed. These runs show bigger weights for the income over cost, the calories, the time spent, the comfort of the trip, and the weather conditions regarding rainy occasions on the contrary of windy conditions. These seem to be reasonable parameters to consider when traveling. The other parameters of the utility function perform "better" in the middle range of the possible interval.

Table 11 aims to summarize the results and bring the most relevant information regarding the parameters.

Table 4
End_time results.

end_time	1200	1230	1260	1290	1320	1350
max	104,13	102,63	105,67	100,75	102,63	104,13
min	92,65	91,07	92,37	94,02	95,33	89,30
mean	96,90	97,31	98,26	96,89	99,25	92,51
range	11,48	11,57	13,30	6,73	7,30	14,83
variance	10,34	9,15	13,55	4,00	8,49	25,20
st_deviation	3,22	3,02	3,68	2,00	2,91	5,02

Table 5
Problem_solved results.

problem_solved	1	2	3	4
max	101,17	385,50	193,41	127,30
min	90,78	359,30	145,73	120,73
mean	96,53	368,88	175,29	124,02
range	10,38	26,20	47,68	6,57
variance	12,23	63,11	343,55	4,89
st_deviation	3,50	7,94	18,53	2,21

Table 3
Start_time results.

start_time	360	390	420	450	480	510	540	570	600
max	101,98	105,67	101,75	100,93	101,72	100,62	104,18	100,82	102,28
min	92,13	91,07	93,27	92,87	91,37	92,65	93,98	93,57	91,07
mean	97,99	97,58	98,00	97,69	96,73	95,62	100,37	97,10	94,25
range	9,85	14,60	8,48	8,07	10,35	7,97	10,20	7,25	11,22
variance	9,58	24,45	9,75	7,82	15,26	5,88	7,24	7,53	11,34
st_deviation	3,10	4,94	3,12	2,80	3,91	2,43	2,69	2,74	3,37

Table 6
Population_size results.

population_size	5	10	15	20	25	30	35	40	45
max	104,2	103,60	104,13	100,87	103,63	100,82	101,42	101,75	101,23
min	89,30	91,37	90,75	90,78	94,05	89,30	93,57	93,57	93,57
mean	96,51	99,41	96,13	96,38	97,95	95,81	97,87	98,70	97,79
range	14,88	12,23	13,38	10,08	9,58	11,52	7,85	8,18	7,67
variance	18,24	11,14	19,02	9,50	9,81	11,60	6,71	6,44	7,69
st_deviation	4,27	3,34	4,36	3,08	3,13	3,41	2,59	2,54	2,77

Table 7
Generations results.

generations	5	10	15	20	25	30	35
max	102,52	101,72	100,82	102,78	101,75	100,82	102,28
min	92,87	91,07	91,10	91,72	91,07	91,37	91,10
mean	97,63	96,60	96,52	97,09	97,13	95,20	96,84
range	9,65	10,65	9,72	11,07	10,68	9,45	11,18
variance	9,12	13,96	14,05	10,82	14,74	10,18	13,61
st_deviation	3,02	3,74	3,75	3,29	3,84	3,19	3,69

Table 8
Cxpb results.

cxpb	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
max	102,6	105,8	103,63	103,63	104,18	100,87	100,82	103,63	101,72
min	91,53	93,27	89,30	89,30	91,37	91,37	89,30	93,27	92,37
mean	97,68	98,06	96,45	98,17	97,33	98,10	96,43	97,31	97,73
range	10,98	12,40	14,33	14,33	12,82	9,50	11,52	10,37	9,35
variance	12,97	14,84	21,46	20,16	13,93	8,14	11,47	12,77	7,61
st_deviation	3,60	3,85	4,63	4,49	3,73	2,85	3,39	3,57	2,76

Table 9
Mutpb results.

mutxpb	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
max	100,8	104,18	102,28	102,28	103,63	104,18	99,43	101,72	101,20
min	91,53	94,23	89,30	91,93	91,53	91,72	89,30	90,75	91,93
mean	95,52	99,35	98,87	97,80	96,01	98,17	95,63	96,57	97,12
range	9,22	9,95	12,98	10,35	12,10	12,47	10,13	10,97	9,27
variance	9,00	10,39	14,79	8,50	12,46	19,16	11,23	11,97	7,86
st_deviation	3,00	3,22	3,85	2,92	3,53	4,38	3,35	3,46	2,80

Table 10
Number_of_alt_locations results.

number_of_alt_locations	1	2	3	4	5	6	7
max	105,67	102,78	103,63	101,17	101,98	101,23	104,13
min	91,37	91,07	91,93	92,13	89,30	94,53	94,05
mean	98,78	98,51	96,66	96,28	96,05	97,57	98,40
range	14,30	11,72	11,70	9,03	12,68	6,70	10,08
variance	15,94	8,34	16,64	9,66	25,59	3,89	8,23
st_deviation	3,99	2,89	4,08	3,11	5,06	1,97	2,87

5. Discussion

Among the papers mentioned in this work, a common search for the optimal values of the main genetic operators is realized. In Table 12, the results demonstrated by various researchers are summarized together with the outcomes of current work.

In their experiments, Beed et al. (Beed et al., 2017) infer that lowering the mutation rate gives more adequate results. On the other hand, increasing the number of generations does not have a huge effect on the outcomes. Moreover, the increase in the population size leads to fast convergence. Unlike other authors, Hassanat et al. (Hassanat et al., 2019) propose a dynamic crossover and mutation rate. The researchers conclude that this approach gives similar results to the commonly used settings: 0,9 crossover rate and 0,03 mutation rate.

Current work has analogous results to other research outcomes showing that a higher crossover rate and population size combined with a lower generation number and mutation rate can provide adequate results (Bakırlı et al., 2011). An important observation is the discrepancy between the population size and the number of generations in this research and other studies (Beed et al., 2017); (Ulukok, 2017) ; (Kinczer and Šulek, 2016) , which is due to the type of the algorithm that provides a heuristic solution. Since the algorithm is built by using low values for the parameters, the range of the testing interval and the step size settings are small, as well. Although these values are smaller than in case of

Table 11
Summary of all results.

Parameter	Summary
start_time	Overall small deviation, but "better" results when the time is higher than 480
end_time	Overall small deviation, but "better" results when the time is higher than 1290
problem_solved	**Car**: Same behavior of default runs **Walking:** Low speed and high impact of route change on time **Public transport**: Inconsistent results related to the huge difference between the travel time of similar routes **Bicycle:** Similar behavior to walking but less impact of route changes
population_size	The more the population, the more consistent the results
generations	Low numbers of generations already give adequate results
cxpb	"Better" results with probabilities above 50 %
mutpb	Adequate results with low and high probabilities, but they can disfigure hereditary characteristics with higher probabilities
number_of_alt_locations	The higher the number of options, the "better" the results
fitnessweight	The inputs that do not show extreme values and the ones with the best performance are described in Annex 2

previous research work, such analysis aiming to optimize activity chains has not been performed yet. Additionally, there is an increased computational cost, which leads to limited space for investigation. Although, the analysis is realized on data collected in Budapest, Hungary, the results can be generally applied to any location. Finally, the number of considered transport modes is limited. The research does not include the newest mobility forms (e.g., e-scooters or shared bikes), yet it covers the most relevant options (car, walking, public transport, and bicycle). The presented results are logical within the context of the algorithm because of the varied population size and the number of generations inside a short range.

The sensitivity analysis method applied in this research provides significant insights into the nature of the parameters. However, the study might be complemented in the future by a global method for "better" quantifying the variability and including the effects from the interaction between the parameters, as indicated by ten Broeke (ten Broeke et al., 2016). According to the results, it can be observed that extreme alternative locations disturb the data and make it difficult to consider the value either as "noise" or not.

Regarding future works, as each transport mode has different behavior and peculiarities, it should be investigated more deeply. This is especially true for public transport since it has some inconsistencies related to the travel time. Regarding the genetic operators, a visible pattern can be seen. The pattern shows a more adequate scenario with a higher crossover rate and population and lower mutation rate and generation, which is a bit different from the applied default values. Therefore, a deeper investigation, which uses a wide range of possibilities reducing the step size and increasing the number of runs, might give a more accurate estimation, as suggested by ten Broeke (ten Broeke et al., 2016).

With the introduction of soft measures through mobile applications, the capacity utilization of transportation systems can be increased without significant infrastructural investments. The essence of the research is to make passenger transportation more efficient by using various data. This study contributes to the optimized use of activity scheduling applications. The behavior of the ACO algorithm is examined under different parameter configurations, which means more adequate results for the users and an increase in the reliability of the applications.

A promising future research direction is the development of an application for users. The application supports the planning of optimized activity chains and the suggestion of sustainable transport modes for urban travelers while considering their personal preferences, as well. Such a complex system using activity planning data and mobility patterns supporting the optimization process and the mode choice at the same time does not exist. Thus, it would be very valuable to elaborate such an application.

From the users' point of view, the application could offer information about the optimal set of activities. Using the optimization algorithm might imply a significant travel time reduction, less emission, and more convenient choices. As a result of the research, personalized recommendations can be provided, which supports the decision-making process during traveling.

The elaborated results are based on the users, who are the main target group of the research, but the outcomes can be beneficial for other stakeholders, too. Thus, supporting the users to make more adequate decisions can have an impact on the transportation system and demand the authorities to adapt to the new patterns. In this way, policymakers can understand which parameters are important to the users and how to deal with the modified travel behavior to increase the quality of the provided service.

6. Conclusion

A daily activity chain can be described as a sequence of activities realized by a person during a day where the activities can be connected through an optimized schedule. The optimization is carried out by the application of the ACO algorithm. As solving this routing problem can require high computational costs, a GA implementation is used, which represents the problem through evolutionary processes. For the investigation of the parameters presented in the algorithm, an OFAT sensitivity analysis is performed. The OFAT can detect the impacts of the

Table 12
A summary of the genetic operators' values.

Author	Population	Generation	Crossover rate	Mutation rate	Remark
Bakırlı et al. (Bakırlı et al., 2011)	250	75	0,9	0,7	The higher the values, the "better" the results
Srinivas et al. (Srinivas et al., 2014)	180	400	0,7 – 0,8	0,4	Fixed population size and generation during the tests
Kinczer et al. (Kinczer and Šulek, 2016)	2000	1000	0,75	0,05 – 0,1	The influence of the mutation rate is bigger than the crossover rate
Beed et al. (Beed et al., 2017)	1000	100	–	0,01	The crossover rate is not analyzed
Ulukok (Ulukok, 2017)	1000	2000	0,7	0,1	Focus on analyzing the population size: the higher, the "better"
Hassanat et al. (Hassanat et al., 2019)	400	1600	Dynamic	Dynamic	Proposed dynamic rates
Daoudi et al. (Daoudi et al., 2019)	150	40—80	0,7	0,001	Investigation of a very low mutation probability
Alamri (Alamri and Effect of Varying the Genetic Algorithm Parameters and Operators on the Optimum PMUs Placement, , 2020)	5000	>=50	>=0,75	<=0,1	Investigation of selection functions
Present work	>=40	5	>=0,5	0,1	A higher number of generations does not have high impact on the results

changes caused by different parameter settings. For conducting the experiment, the relevant parameters connected to activity chains, transport modes, optimization algorithms, and fitness functions are identified. A parameter optimization framework capable of running all the tests consecutively is developed.

The primary aim of the paper is to perform a detailed parameter analysis and present the most suitable settings for a specific transportation problem. The results of this work provide relevant information on how the proposed algorithm works with a different set of parameters. Additionally, the research demonstrates that complementing approaches are needed to better understand the functioning and the calibration of the ACO system. With a properly elaborated application, users can utilize the benefits of the optimization algorithm. Optimized routes result in a significant travel time reduction and contribute to less emission when planning daily activities in the field of transportation.

CRediT authorship contribution statement

Matheus Moro Zamprogno: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Domokos Esztergár-Kiss:** Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was supported by the János Bolyai Research Fellowship of the Hungarian Academy of Sciences (BO/00090/21/6). The linguistic revision of current paper is prepared by Eszter Tóth.

Annex.

Annex 1: Testing intervals.

Parameter	A	B	C	D	E	F	G	H	I
start_time	360	390	420	450	480	510	540	570	600
end_time	1200	1230	1260	1290	1320	1350			
problem_solved	1	2	3	4					
population_size	5	10	15	20	25	30	35	40	45
generations	5	10	15	20	25	30	35		
cxpb	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
mutpb	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
number_of_alt_locations	1	2	3	4	5	6	7		
fitnessweight1	0	1	2	3	4	5	6	7	
fitnessweight2	0	1	2	3	4	5	6	7	
fitnessweight3	0	1	2	3	4	5	6	7	
fitnessweight4	0	1	2	3	4	5	6	7	
fitnessweight5	0	1	2	3	4	5	6	7	
fitnessweight6	0	1	2	3	4	5	6	7	
fitnessweight7	0	1	2	3	4	5	6	7	
fitnessweight8	0	1	2	3	4	5	6	7	
fitnessweight9	0	1	2	3	4	5	6	7	
fitnessweight10	0	1	2	3	4	5	6	7	
fitnessweight11	0	1	2	3	4	5	6	7	
fitnessweight12	0	1	2	3	4	5	6	7	
fitnessweight13	0	1	2	3	4	5	6	7	
fitnessweight14	0	1	2	3	4	5	6	7	
fitnessweight15	0	1	2	3	4	5	6	7	
fitnessweight16	0	1	2	3	4	5	6	7	

Annex 2: Summary of the fitness weight results.

Parameter	Default value	Best performance
fitnessweight1	0	6
fitnessweight2	0	3
fitnessweight3	0	7
fitnessweight4	4	7
fitnessweight5	0	0
fitnessweight6	4	3
fitnessweight7	4	7
fitnessweight8	4	3
fitnessweight9	4	4
fitnessweight10	7	5
fitnessweight11	0	5
fitnessweight12	0	6
fitnessweight13	0	5
fitnessweight14	7	2
fitnessweight15	0	1
fitnessweight16	4	4

References

B. Alamri, Effect of Varying the Genetic Algorithm Parameters and Operators on the Optimum PMUs Placement, 2020 55th International Universities Power Engineering Conference, UPEC 2020, pp. 1–6, doi: 10.1109/UPEC49904.2020.9209812.

Bakırlı, G., Birant, D., Kut, A., 2011. An incremental genetic algorithm for classification and sensitivity analysis of its parameters. Expert Systems with Applications 38, 2609–2620. https://doi.org/10.1016/j.eswa.2010.08.051.

Bast, H., et al., 2016. Route planning in transportation networks. Lecture Notes in Computer Science 9220, 19–80. https://doi.org/10.1007/978-3-319-49487-6_2.

R. S. Beed, S. Sarkar, A. Roy, and S. Chatterjee, **A Study of the Genetic Algorithm Parameters for Solving Multi-objective Travelling Salesman Problem**, *2017 International Conference on Information Technology, ICIT 2017*, pp. 23–29, 2018, doi: 10.1109/ICIT.2017.49.

Camargo Pérez, J., Carrillo, M.H., Montoya-Torres, J.R., 2015. Multi-criteria approaches for urban passenger transport systems: a literature review. Annals of Operations Research 226 (1), 69–87. https://doi.org/10.1007/s10479-014-1681-8.

Charypar, D., Nagel, K., 2005. Generating complete all-day activity plans with genetic algorithms. Transportation 32 (4), 369–397. https://doi.org/10.1007/s11116-004-8287-y.

Daoudi, K., Boudi, E., Abdellah, M., 2019. Genetic approach for multiobjective optimization of epicyclical gear train. Mathematical Problems in Engineering 2019, 9324903. https://doi.org/10.1155/2019/9324903.

S. Dass and A. S. Namin, **A sensitivity analysis of evolutionary algorithms in generating secure configurations***, 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 2065-2072, 2020, doi: https://doi.org/10.1109/BigData50022.2020.9378307.

Esztergár-Kiss, D., 2020. **Trip Chaining Model with Classification and Optimization Parameters**, *Sustainability (switzerland)* vol. 12, 16. https://doi.org/10.3390/SU12166422.

D. Esztergár-Kiss and Z. Rózsa, **Simulation results for a daily activity chain optimization method**, *2015 International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2015*, pp. 259–264, 2015, doi: 10.1109/MTITS.2015.7223265.

Esztergár-Kiss, D., Remeli, V., 2019. Toward practical algorithms for activity chain optimization. Transportation Letters 13 (1), pp. https://doi.org/10.1080/19427867.2019.1702250.

Esztergár-Kiss, D., Rózsa, Z., Tettamanti, T., 2018. Extensions of the Activity Chain Optimization Method. Journal of Urban Technology 25 (2), 125–142. https://doi.org/10.1080/10630732.2017.1407998.

Esztergár-Kiss, D., Rózsa, Z., Tettamanti, T., 2020. An activity chain optimization method with comparison of test cases for different transportation modes. Transportmetrica a: Transport Science 16 (2), 293–315. https://doi.org/10.1080/23249935.2019.1692958.

Fan, W., Machemehl, R.B., 2006. Optimal Transit Route Network Design Problem with Variable Transit Demand: Genetic Algorithm Approach. Journal of Transportation Engineering 1, 40–51. https://doi.org/10.1061/(ASCE)0733-947X(2006)132:1(40).

Fazayeli, S., Eydi, A., Kamalabadi, I.N., 2018. Location-routing problem in multimodal transportation network with time windows and fuzzy demands: Presenting a two-part genetic algorithm. Computers & Industrial Engineering 119, 233–246. https://doi.org/10.1016/j.cie.2018.03.041.

Hassanat, A., Almohammadi, K., Alkafaween, E., Abunawas, E., Hammouri, A., Prasath, V.B.S., 2019. Choosing mutation and crossover ratios for genetic algorithms-a review with a new dynamic approach. Information (switzerland) 10 (12), pp. https://doi.org/10.3390/info10120390.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor.

Kinczer, T., Šulek, P., 2016. The impact of genetic algorithm parameters on the optimization of hydro-thermal coordination. Pollack Periodica 11 (2), 113–123. https://doi.org/10.1556/606.2016.11.2.10.

Sinha, K.C., 2003. Sustainability and urban public transportation. Journal of Transportation Engineering 129 (4), 331–341. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:4(331).

Srinivas, C., Reddy, B.R., Ramji, K., Naveen, R., 2014. Sensitivity Analysis to Determine the Parameters of Genetic Algorithm for Machine Layout. Procedia Materials Science 6, 866–876. https://doi.org/10.1016/j.mspro.2014.07.104.

ten Broeke, G., van Voorn, G., Ligtenberg, A., 2016. Which sensitivity analysis method should i use for my agent-based model? Jasss 19 (1), 1–36. https://doi.org/10.18564/jasss.2857.

M. K. Ulukok, **Impact of genetic algorithm's parameters on solution of numerical optimization benchmark problems**, *9th International Conference on Computational Intelligence and Communication Networks, CICN 2017*, vol. 2018, pp. 139–143, 2018, doi: 10.1109/CICN.2017.8319373.

Wirsansky, E., 2020. Hands-On Genetic Algorithms with Python, 1st ed. Packt, Birmingham.

Yang, B., Hu, Z.-H., Wei, C., Li, S.-Q., Zhao, L., Jia, S., 2015. Routing with time-windows for multiple environmental vehicle types. Computers & Industrial Engineering 89, 150–161. https://doi.org/10.1016/j.cie.2015.02.001.

Yin, L., Lin, N., Zhao, Z., 2021. Mining Daily Activity Chains from Large-Scale Mobile Phone Location Data. Cities 109, 103013. https://doi.org/10.1016/j.cities.2020.103013.