

Warszawa, October 3rd, 2024

dr hab. inż. Karol Opara, prof. IBS PAN
Instytut Badań Systemowych PAN
ul. Newelska 6
01-447 Warszawa

Report for the PhD thesis
“Designing Urban Transit Network using Memetic Algorithm”
by Hanan Ali Mohammed Ba Ali

Thesis scope and structure

Urban transit network design is a prominent topic of practical significance. The thesis is dedicated to its strategic level, understood as finding a possibly good set of bus and tram routes in a city without timetable optimization or operational planning. Chapter 1 presents the context of transit network design. This task can be addressed with metaheuristics, briefly surveyed in Chapter 2. In Chapter 3, the Author introduces a memetic algorithm, which makes the most substantial contribution of this thesis. The proposed method comprises a variant of the genetic algorithm refined by several route exchanges called hill climbing local search. This algorithm was run on four literature networks and proved comparable to or even better than the literature algorithms. It was also applied to the Kraków tram network. Chapter 4 reinvents graph smoothing and graph generation via a configuration model. It also integrates these methods to create additional networks resembling the Kraków one. Chapter 5 addresses the need to modify the existing urban transit network gradually. The thesis is summarized in Chapter 6.

General assessment

The thesis lacks significant novelty. Original ideas concentrate predominantly in Chapter 3 and focus on the design and validation of the memetic algorithm. Although the proposed techniques seem natural, their integration and successful application to a few networks bring certain technological merit.

Chapter 4 largely rediscovers classical methods of network theory. This is unnecessary and suggests that the Author might have studied that field insufficiently. Surrogate modelling utilizing graph compression is appropriate, yet its benefits have not been verified for large networks.

Chapter 5 outlines a promising and necessary idea to re-design the urban network through gradual modifications. In my opinion, the proposed approach is unlikely to satisfy either the transportation policy decision-makers or the end users of the network. The overall merit of the thesis would benefit from proposing a more practical transformation plan, which seems doable.

The long list of detailed remarks provided in the following section demonstrates a variety of issues with the thesis concerning algorithmics, methodology of simulation experiments, adequacy for practical urban transit network design, and presentation of the results. Although most of these issues are fixable, their prevalence impacts the overall quality of the work.

Detailed remarks

The list below presents detailed remarks, starting with the major ones about algorithmic aspects, the evaluation methodology and simulation experiments.

1. Linearization of a three-criterial optimization problem (3.5) is a common approach, yet it has certain limitations. In particular, in the case of concave Pareto fronts, some of the non-dominated solutions cannot be achieved for any combination of weights. A geometric mean could be a better option. It would also alleviate another issue, namely unit incoherence. The first term of the sum (3.5) is in unstated temporal units, probably minutes, whereas the second and third are percentages. This issue should be addressed explicitly, as otherwise, representing the same optimization problem in different units, say seconds, would implicitly modify the linearization weights.
2. It is misleading to present the computational budget in terms of the number of generations, as the time for finalizing an iteration varies for different algorithms. This is clearly visible in Table 3.8, where the sum of wall clock times for the genetic algorithm and hill climbing local search equal the time for the proposed memetic algorithm up to a rounding error of 0.02 seconds for all considered cases. It is widely accepted to compare metaheuristics for the same number of fitness function evaluations. I wonder why this metric was not utilized here and whether the superiority of the proposed memetic algorithm over the genetic algorithm and hill climbing local search presented in Figure 3.18 would hold if the computational budgets for the latter two methods were appropriately increased (from 30% to 70% depending on a case).
3. Chapter 4 focuses on a graph compression technique that “operates by consolidating consecutive nodes’ data that possess only one or two edges into a singular node”. I wonder why so much space was dedicated to such a basic idea. Moreover, it is well-known under the name of graph smoothing¹. Hence, most of section 4.1.3 could be described as looking for the topological graphs with a minimal number of nodes. What is the novelty of this approach? This section also contains simple mathematical derivations with minor issues, e.g., equation (4.1) should have implications from right to left rather than the equivalence (iff) operator.
4. Listing 4.4 closely resembles the configuration model² for generating random networks from a given degree sequence. The author's presentation of the results would have been clearer if she had explicitly referred to the background knowledge and focused on her original results. I also wonder if any precautions were taken to avoid generating graphs with disconnected components that seem unnatural for urban transit networks.
5. The Author stops her considerations somewhat early after solving the transit problem for a compressed graph. I wonder why this solution was not translated back into the original graph? One could even run a few more iterations of the local search to fine-tune the solution obtained using the compressed graph as a surrogate model.

¹ See <https://mathworld.wolfram.com/GraphSmoothing.html> and references therein

² See https://en.wikipedia.org/wiki/Configuration_model

6. Chapter 4 begins with a statement that “graph compression is the process of reducing the storage space required to represent a graph (...)”. However, in Listing 4.8 possibly huge travel time matrix is utilized in recursively. Matlab passes matrices by value. Hence, for larger networks, execution of function from Listing 4.8 will presumably use up all the available memory. I wish the Author would select a more memory-efficient implementation by removing the recurrence or using sparse matrices.
7. Random graphs obtained by reconfiguring the nodes of the Kraków network can be compressed by approximately 50%. This improves the execution time of the memetic algorithm by up to 60%. For a complete picture, one should account for the graph compression time as an overhead in the computation of this improvement. More importantly, though, adding a whole compression process to save half an hour of execution time sounds peculiar. Note that the urban design problem is addressed only occasionally, presumably not more often than once a year.
8. The Author repetitively motivates the need for Her solutions by challenges related to planning transit in “large networks with hundreds of transit vehicle routes and thousands of stops”. The discussion suggests success, e.g., “genetic algorithm can effectively explore the huge search space of the urban transit network design”. However, all networks analysed in the thesis have less than 150 nodes. Surprisingly, this relatively small size concerns also artificially generated networks. It is a pity that the Author did not use this opportunity to demonstrate the scaling of Her algorithms and the benefits of the proposed graph compression technique. A convincing example based on the Kraków city network could take over 1000 bus stops as nodes, extract the possible links between them using cartographic resources (OpenStreetMap, Geoport, municipal resources, etc.) and apply the proposed methods. This would show, how to redesign the currently utilized 71 standard bus routes together with 71 agglomerative bus routes and 56 other routes (including 21 night busses).
9. I appreciate the idea described in Chapter 5, namely proposing gradual changes from the current route design to the newly optimized one. However, its operationalization is less than perfect. Does the Author really expect the users of the transit network to temporarily accept much worse solutions than the baseline, as Figure 5.5 seems to suggest? I think it would be more politically and socially acceptable to aggregate consecutive steps into a larger one if they deteriorate the end user's experience much. I also wonder if the Author has considered adding a regularization term penalizing for the discrepancy from the previous solution to directly create a path of improvements.

Issues related to urban transit planning

10. The assumption that each solution has a fixed number of routes (section 3.1.1) requires more rationale. Is it just an issue of technical convenience or a more fundamental requirement arising in transportation planning?
11. The “Kraków transit network”, discussed in the thesis, refers to the tram line (although it was not immediately apparent to me as a reader). Part of the reason was the unintuitive visualization of the Kraków transfer network in the thesis, which keeps the topology but greatly distorts spatial information, as shown in Fig. 1 of this review. Presenting the transit network as a map would make the results much more easily understandable to readers familiar with the Kraków city transit system and transportation professionals.

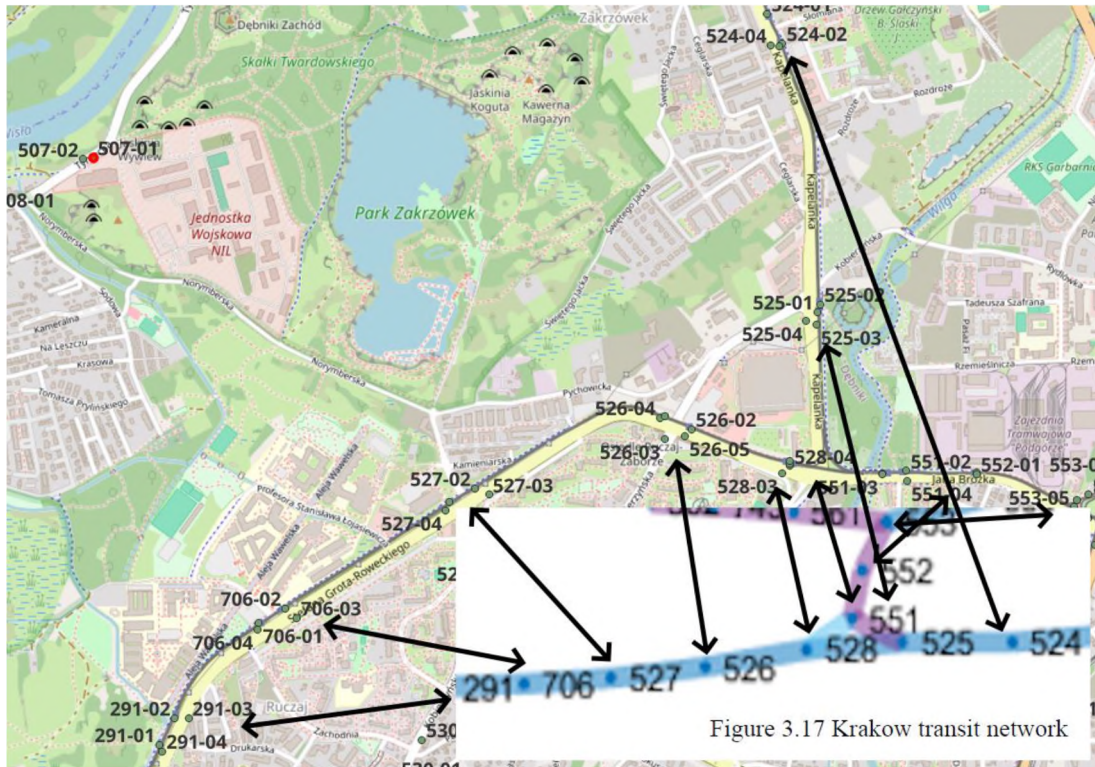


Fig. 1. Comparison of the visualization of the tram network from the thesis and the corresponding Kraków city data overlaid on OpenStreetMap

12. Trams cannot make a U-turn without appropriate infrastructure. Currently, this aspect seems unaddressed. Therefore, shortening the route that exceeds the length constraint may create infeasible solutions.

Issues related to presentation of the thesis

13. The thesis contains multiple long listings of Matlab code. It is advisable to present algorithms in pseudocodes using standard mathematical notation. Current and future readers may not be familiar with the particular programming language. For instance, it is non-obvious that function `sum(X, 2)` computes row sums of matrix `X`. Listings could be added to an appendix. Ideally, the code, exemplary data and documentation should be available online.

14. The structure of the thesis is somewhat unusual. It begins with an “Introduction” section, which resembles an executive summary that provides the final results with percentage improvement on the first page. The following chapters intertwine elements of the background and literature presentation, methods and results. I believe that a stricter separation of those parts would improve the presentation of the thesis.

15. The parts of the thesis that describe the original contribution and code listings are very technical and look more like a software manual than a PhD thesis. Sometimes, they are also too basic for a knowledgeable reader. For instance, Listing 3.10 and the following paragraph detail a method for checking if a vector has repeated elements. Similarly, Listing 4.1 checks whether percentage vectors sum to 100. Moreover, it appears to test the equivalence of floating point numbers using a negated equality operator, which is prone to numerical issues. For example, in Matlab $1/2 + 1/3 + 1/6 - 1$ is not zero but $-1.1102e-16$. Below is another example of an excessively detailed and technical style (page 117):

Therefore, we use the function *find* to get the indices of all node neighbors where *inf* refers to unconnected nodes in our matrix *fullTravelTimeMatrix* and the rest are connected. This process includes checking each node in our network as *nd* variable as in equation (4.6)

$$[indNeigh] = find(\sim isinf(fullTravelTimeMatrix(nd, :))) \quad (4.6)$$

In the next step, the function *getOrderOfNode* is used to get the number of nodes that are connected to an input node as shown below in Listing 4.6.

On page 152, the relation clearly presented with arrows in Figure 5.4 is tediously repeated, which is particularly surprising since no further details about these routes are provided:

The rest of the routes are assigned to each other based on the HA as follows: current Route 2 should be replaced with proposed Route 3, current Route 3 with proposed Route 4, current Route 4 with proposed Route 20, current Route 6 with proposed Route 11, current Route 8 with proposed Route 15, current Route 13 with proposed Route 18, current Route 17 with proposed Route 6, current Route 18 with proposed Route 22, current Route 19 with proposed Route 7, current Route 20 with proposed Route 2, current Route 22 with proposed Route 13, current Route 49 with proposed Route 16, and current Route 52 with proposed Route 14.

16. There are some inconsistencies between different parts of the thesis. For example, Table 3.1 states that the Kraków network has 149 nodes and 535 links. The average of 3.6 links per node seems implausible when contrasted with the network graph in Figure 3.15. Furthermore, Table 4.2 suggests that the total sum of out-degrees is only 328.

Minor issues. Addressing them is likely to improve the overall quality of the thesis.

17. Multiple editorial aspects need improvement.

- a. Paragraphs should be visually separated by vertical space or an indent. Currently, this is not the case, as is evident on page 13.
- b. Word capitalization is incorrect (“Furthermore, Transit agencies...”) and analogously on pages 94, 155, and 158.
- c. On pages 135, 140, and 164 (twice) the whole sentences are repeated in consecutive paragraphs with hardly any change.
- d. Occasionally, there are typos, such as a double colon on page 23 or an incorrect index in Conjecture 1, where the last symbol should be $f(S_2')$.
- e. Surprisingly, Conjecture 1, shown in the thesis to be false for certain functions f , is presented four times, on pages 24, 109, 129 (renamed to Property 1), and 162.
- f. Mathematical formulas are a bit messy. For example, the superscript in the crowding distance variable cd^i , occurring in equation (2.1), unexpectedly changes to “cd_i”.

18. There are a few issues related to the listings:

- a. Can the for loop in Listing 3.3 make more than a single iteration?
- b. In Listing 3.5, do you support the case when two or more routes go through a certain sub-path? For instance, two different routes on the same bridge over a river?
- c. There are occasional syntax errors, such as missing “end” in Listing 3.8 or non-uniform notation in the for loops in Listing 4.7.
- d. Some variables are italicised while others are not (Listing 3.9).

19. On page 21, we read that the network design problem is NP-hard. In the following paragraph, it is stated that “MA [memetic algorithm] (...) is used to find the best possible structure that

represents the flawless transit network”. Further discussion makes it clear the Author is aware that Her metaheuristics do not guarantee the optimality of the solution. However, the current phrase is unfortunate as it may mislead readers inexperienced in metaheuristic research.

20. There are some disputable claims. For example, on page 22, the Author states “The solution space for the UTND in large areas is too large and complex to be solved with traditional mathematical optimization techniques...”. I accept the complexity argument, but if it is indeed too large for mathematical optimization (of which I am not so certain, given recent progress in the mixed integer programming solvers) its size would make it even more prohibitively large for metaheuristics.
21. Figure 2.2 is claimed to present a “two-dimensional solution space” whereas the depicted solution space is one-dimensional.
22. The description of the NSGA-II algorithm tailored by Chai and Liang (2020) “to address the Transit Network Design and Frequency Setting Problem” misses two crucial pieces of information: that it solves the multicriterial optimization problem and the modifications by Chai and Liang (2020) against the classical formulation of NSGA-II.
23. Section 2.2.3 describes the Greedy Randomized Adaptive Search Procedure (GRASP) but does not mention any modifications related to urban transit planning. GRASP is also absent from the literature overview in Table 2.1 and not referred to later in the thesis. What was the point of adding this section?
24. “Hill Climbing Local Search” is not the most suitable name for an algorithm operating in a discrete space without gradients. There seem to be more fitting names, built around the notions such as route exchange and greediness.
25. A broader discussion of the rationale behind setting the selected values of weights would be useful. The same considers different scenarios and parameters used in simulations.
26. The statement on page 71 that “rank selection is computationally expensive due to the sorting requirement” is surprising as the log-linear sorting cost is typically negligible in comparison with the cost of the fitness function evaluation. To assess the algorithmic overhead it would be interesting to learn the percentage of time the proposed algorithm spends on evaluating the fitness function. The following sentence is also disputable: “the roulette wheel method (...) may be perceived as unfair because the worst solutions never become parents”. First, genetic algorithms are about optimization rather than “fairness”. More importantly, in standard roulette wheel selection, even the worst individual has some probability of becoming a parent.
27. On page 50, we correctly read that “elitist solution is a deterministic selection method”. However, on page 73 we are surprised by a statement that “The fitDistribution vector contains solution indices along with their fitness values' probability distribution, computed using the formula (3.5) to ensure that the elite solution has a significantly higher chance of being selected.” What “chance” is the Author referring to, granted that we are considering a deterministic method? Moreover, the “fitness function” formula (3.5) does not provide a probability distribution, which is illustrated by the fact that the numerical values of this formula presented in the thesis are usually negative.
28. In the urban network design problem, there is a requirement that routes cannot have loops (page 21). However, quite surprisingly, Route 4 in the first member of the memetic population shown in Figure 3.11 seems to have a loop ($8 \rightarrow 5 \rightarrow 4 \rightarrow 5$) despite “consistently verifying” and “disregarding invalid solutions” (page 59). In practice, most reasonable transit routes would indeed avoid loops. Nevertheless, excluding such cases arbitrarily requires some justification. For example, the *Ringstrasse* tram in Vienna cycles around the city centre.

29. The travel time dataset is assumed to be constant, i.e., it does not account for traffic jams during peak hours. This limitation should be stated and discussed explicitly. It seems justified for trams, but buses are slowed down by high traffic unless they have a dedicated lane. Next, the symmetry of the time travel matrix is disputable for urban road traffic due to daily commuting patterns.
30. Finding a solution that improves all four criteria for the Kraków tram network is a promising outcome from the memetic algorithm. Mandl network is so small that with eight or seven routes it was possible to “achieve 100% of the total transfer demands satisfied directly”. For Mumfold networks, the proposed algorithm outperforms the competition in terms of direct routes at a cost of longer average travel times. Unfortunately, despite noticing this trade-off, the Author does not explore it. It would be interesting to see if the proposed memetic algorithm for different settings of weights can find a solution that dominates literature results for all the criteria.
31. It is confusing that the fitness function, defined in equation (3.5), takes the lowest value for the best, “most fit” solutions. Moreover, the numerical values of this fitness, shown, e.g., in Figure 3.12, are negative, which also puzzles me about how the roulette wheel selection was implemented for such a case.
32. Table 3.6 compares the proposed memetic algorithm against the literature results. In its description, we read that “the obtained results are significantly better than other approaches”, even though no statistical test was used to justify the claims about significance. In comparing metaheuristic algorithms, it is recommended to type in bold the best algorithm for each test problem and those that do not perform statistically significantly worse. This also allows for aggregating the results across the whole benchmark by computing fractions of tasks in which the proposed method was among the best algorithms.
33. Comparisons against other algorithms, whose results are presented in Tables 3.6 and 3.7, show that the best out of multiple (20 or 100) runs of the proposed memetic algorithm is competitive and often outperforms the competitors. However, the mean performance is notably weaker. Hence, two questions arise. What are the computational budgets of the literature algorithms, and are they comparable to the budget utilized in the proposed method? Has the Author considered restarting multiple times or running in many parallel instances to find the best solution as the default approach?
34. The average execution time of the proposed memetic algorithm is less than 5 minutes for Mandl networks and less than 25 minutes for Mumford networks (Table 3.8). Why so short? Do the algorithms robustly converge to the optimum within this time? Urban transit network design can take weeks or months and is not repeated too often. Potentially, in this problem, the computational budget is as large as one wishes and the infrastructure allows. In particular, one can leave the computations running overnight or during the weekend for several hours without disturbing their work. Accessing clusters is also easy, for instance, through PLGrid infrastructure.
35. Figure 3.16 suggests that the average fitness value gradually improves with the number of routes increasing from 19 to 24. I wonder why an even larger number of routes was not checked? Maybe the optimum could be achieved for 25, 30 or 60 routes?
36. On page 91, it is stated that “we computed the fitness values of the pure GA and MA with mutation operator disabled (just HCLS) in Mandl_4 routes and Mumford3 to evaluate the efficiency of our suggested algorithm (MA). The experiments on Mandl and Mumford networks show that the suggested MA provides significant enhancements with respect to the values of the fitness function (...)”. Why was this experiment done only using an arbitrary subset of networks? Do the results hold also for other variants of Mandl and Mumford networks?

37. I wonder why Figures 3.18 and 3.19 use two different normalization ranges. The non-standard range from 0 to 20 seems especially puzzling.
38. On page 92, we read “The experiments illustrate that MA consistently progressed throughout the iterations. In contrast, GA showcased gradual progress, benefitting from the mutation operator to prevent being trapped in local minima states, as evident in Figures 3.18 and 3.19. However, HCLS reached a point where it couldn't generate better route combinations than the existing ones and without the mutation operator, this led to a scenario of local minima status.” I do not see the contradiction between the first two sentences implied by connector “in contrast”. Moreover, the convergence curves suggest that the proposed memetic algorithm might have also stagnated at a local optimum, albeit better than the one found by the hill climbing local search.
39. The first paragraph in section 4.1.1.2 needs a citation and paraphrasing. The statement “a base model is employed to produce a predictive posterior distribution for each node (...)” provided without a wider context is more likely to occlude than clarify the described approach.
40. In Table 4.1 and Figure 4.1, the time is, somewhat surprisingly, treated as an integer from 1 to 6 except 5. The unit is unstated. I probably can guess the reasons for such a choice, but they should be stated explicitly.
41. It is unclear what “theoretical demand” refers to in Figure 4.3. Moreover, providing the demand unit would facilitate understanding of that graph.
42. Randomized demand is a crude approximation of daily migration patterns in a city. It would be more adequate to reflect commuting work in the city centre, e.g., by increasing the transportation demands to and from nodes with high centrality.
43. I find it surprising that instead of assigning the total travel time over a path for its compressed representation, the Author proposes to use minimum, average or maximum travel time. The difference between these cases can be notable if many nodes are compressed. This is sometimes the case, e.g., the rightmost branch of Graph A in Figure 4.7 contains 9 nodes that are collapsed into a single one. I struggle to imagine viable transportation planning questions that could be addressed by the minimum or maximum travel time between consecutive stops on tram line section forming a “simple path”.
44. Conjecture 1 is argued to hold when the fitness function is defined as the number of transits. This is only true if the graphs do not contain a “simple path” longer than the maximum route length.
45. On page 131, there is an imprecise statement “Figure 4.8 shows that there is a very strong positive correlation between our data sets and rejects the null hypothesis that there is no correlation”. Although these datasets are indeed clearly correlated, alluring that the statistical test was based on a figure is a misuse of mathematical formalism.
46. I would not call disruptions in the correlation coefficient from 1 to 0.63 and 0.83 “relatively minimal” (page 133).
47. There are numerous issues with Figure 5.5. The horizontal axis, labelled “cost” is non-monotonous, yet the plot shows a smooth curve. Next, for “cost” equal to 144, we have two different fitness values (around -5 and -22). Maybe the horizontal axis shows consecutive steps? For such discrete data, one should use disconnected points rather than a smooth curve.
48. Chapter 6 summarizes the thesis, yet it mentions a new element of the optimization problem: a constraint on fleet capacity. Was it used at all in the considered model?
49. Typos: Page 75. “80% of the solutions are selected by using a crossover operator” → “80% of the solutions are created by the crossover operator”, “other edges and vertex” → “other edges

and vertices”, “simplifies the original graph size as illustrated in the example below” →
“reduces the original graph size as illustrated in the definition below”.

Conclusion

The thesis aims to explore a relevant and complicated problem of urban transit network design. Unfortunately, the novelty of the proposed ideas is limited. Moreover, there are numerous problems with the substantive and presentation aspects of the research. While the effort to contribute to the field is acknowledged, the thesis fails to demonstrate the scientific soundness necessary to meet the rigorous standards required for a PhD. Hence, I cannot recommend awarding this degree. Significant revisions focusing on algorithmic ideas, simulation experiments, and presentation would be needed to enhance the quality and impact of this work.



Karol Opara