Evolving Robust Facility Placements

William Thompson*
Alex Friedrichsen*
williamthompson@uvm.edu
alexanderfriedrichsen@uvm.edu
University of Vermont
Burlington, Vermont, USA

ABSTRACT

Climate change is a growing issue. Placement locations of facilities are a constantly relevant question for hospitals, restaurants, post-offices, and any number of other facility types. Access to facilities should be equitable based on population density, and the minimum distance to the facility is desired to be minimizes.

CCS CONCEPTS

• **Computing methodologies → Genetic algorithms**.

KEYWORDS

P-median problem, Genetic Algorithms, Julia Language, Robustness, Evolutionary Algorithms, Spatial Distributions, Location-allocation

ACM Reference Format:
William Thompson and Alex Friedrichsen. 2021. Evolving Robust Facility Placements. In *Proceedings of the Genetic and Evolutionary Computation Conference 2021 (GECCO '21).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In this paper, we present a study on the application of an evolutionary algorithm to the p-median problem, a class of facility location problems in which the aim is to identify p source locations and map them to n destinations while minimizing the average distance between destinations and sources. Our algorithm, which is implemented in the Julia programming language, was used to solve the p-median problem on a map of the United States using census data on population distribution.

Each genome is represented as a list of coordinates corresponding to the locations of the facilities. The algorithm uses N-point crossover to combine genomes from the parent population to create offspring, with a small probability of mutation in which a random facility is moved by a small distance. Elitism is implemented to ensure that the best genomes from the parent population are preserved in the next generation, and tournament selection is used to select the parents for crossover and mutation.

We evolved the placements of 500 facilities over 1000 generations and ran the algorithm multiple times to explore the effect of different parameter combinations. We found that the algorithm was able to evolve placements that followed the expected two-thirds scaling relationship for the optimal solution.

In addition to optimizing the placement of facilities, we also investigated the ability of the evolutionary algorithm to adapt to perturbations that occur during the optimization process. To simulate catastrophic events that destroy facilities, we introduced perturbations into the algorithm that randomly destroy facilities within a given radius and calculate the fitness of the new configuration. We then compared the performance of the evolved placements with and without perturbations to assess their robustness to these events. Our results show that the evolutionary algorithm is able to evolve placements that are more robust to perturbations, with a lower decrease in fitness relative to the non-evolved placements.

In the following sections, we will describe the p-median problem and existing approaches to solving it, provide details of the evolutionary algorithm and perturbation technique used in our study, present the results of our experiments, and discuss the implications of our findings

2 RELATED WORK

Previous research has shown that there is a positive relationship between population density and the density of facilities such as grocery stores, schools, and fire stations. This relationship can be described by a power law, with the facility density being proportional to the population density to the power of 2/3 [3] [4]. While our research does not currently take facility type into account, empirical analysis has found that the power law exponent can vary depending on the type of facility. A model that incorporates economic mechanisms has been proposed to explain this variation and shows that commercial facilities have an exponent of 1, while public facilities have an exponent of 2/3. These results align with empirical data.

2.1 Evolutionary dynamics and highly optimized tolerance

Zhou *et al.* [5] present a numerical model for the evolution of complexity in a lattice community based on the concept of Highly Optimized Tolerance (HOT), which emphasizes the robustness tradeoffs that dominate engineering design and, according to the authors, biological evolution. The model is used to explore scenarios for the evolution and extinction of organisms in different environments, including the effects of different habitats and mutation rates on the

*Both authors contributed equally to this research.

GECCO '21, July 10–14, 2021, Lille, France

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

adaptation and diversity of the organisms, as well as the competition between generalists and specialists. The model is based on engineering concepts and uses methods from statistical physics to illustrate the robustness tradeoffs that are central to the model. The results show that the model exhibits a range of microevolutionary and macroevolutionary phenomena, including the emergence of generalists and specialists in different habitats, the role of fast and slow mutators in adaptation, and the phenomenon of punctuated equilibrium. The application of the HOT model is relevant to our work as another study of robustness behavior.

Carlson and Doyle (coauthors on the paper mentioned above) contrast the emphasis on design and rare configurations in the Highly Optimized Tolerance (HOT) model with the perspective provided by NSOC/CAS/SOC, which emphasizes structural complexity as "emerging between order and disorder" at a bifurcation or phase transition in a random interconnection of components [2]. This approach, which is inspired by critical phenomena and other concepts from statistical physics and dynamical systems, suggests that the details of component behavior and interconnection are largely irrelevant to system-wide behavior. This perspective has implications for the study of facility placement and other urban planning problems

Alghanmi *et al.* [1] present an evaluation of existing heuristic algorithms for the p-median problem, a class of facility location problems in which the aim is to identify p source locations and map them to n destinations while minimizing the average distance between destinations and sources. The study adds to the current literature by providing a thorough evaluation of classic heuristics and investigating the effect of the spatial distribution of destinations and the number of sources and destinations on the performance of the algorithms for varying problem sizes using synthetic and real datasets. The performance of the algorithms is evaluated using the objective function value, execution time, and stability of the solution. The sensitivity of the algorithms to the spatial distribution of destinations and the scale of the problem is analyzed, and the utility of the study is demonstrated by applying the algorithms to the selection of locations for ad-hoc clinics in a bio-emergency scenario. The results show that interchange algorithms perform well in terms of execution time and cost function values and are more stable for clustered distributions.

3 METHODOLOGY

In this section, we describe the methodology and experimental setup used in our study on the application of an evolutionary algorithm to the p-median problem. We first provide details on the evolutionary algorithm we implement, followed by parameter selection for the evolutionary algorithm, followed by a description of the parameter sweep and tuning used to explore the performance of the algorithm for different parameter combinations. We then describe the p-median optimization null model used as a baseline for comparison, and finally, we describe the implementation of the catastrophe perturbations used to assess the robustness of the evolved solutions.

3.1 Evolutionary Algorithm

Our evolutionary algorithm was implemented in the Julia programming language and uses a mu + lambda evolutionary algorithm, in which a population of potential solutions (genomes) is evolved over multiple generations. Each genome is represented as a list of coordinates corresponding to the locations of the facilities within the contiguous US. A population of 30,000 "citizens" are placed on the map. Each county within the US is allotted a number of citizens proportionate to the percentage of the US population residing within that county. Each counties allotment of citizens are distributed uniformly within the borders of the county.

$$F = -\frac{1}{N}\sum_{i=1}^{p}\sum_{j}^{N}\min_{i\in\{1...p\}} \mid \mathbf{r_j} - \mathbf{r_i} \mid$$

Where $\mathbf{r}_i \in \{\mathbf{r}_1...\mathbf{r}_p\}$ are the locations of facilities and $\mathbf{r}_j \in \{\mathbf{r}_1...\mathbf{r}_M\}$ are the locations of the citizens within the population. This is equivalent to the average distance traveled to the nearest facility. A nearest neighbor search is used to ensure that distances are only evaluated between citizens and the nearest facility. The mutation operator moves N facilities in the genome to random new locations within the contiguous US. We include an N point crossover operator in our evolutionary algorithm. For crossover, two parent genomes are chosen and a set of N indices are chosen, the parent genomes are split into subsections. These indices are child genomes are created by swapping alternate subsections of the parent genomes.

For our selection procedure, we implement a tournament selection. N individuals are randomly included in a tournament and from this set the M individuals with the highest fitness are chosen as tournament winners. The tournament winners are inserted into the population for the next generation. This pattern is repeated until μ individuals have been selected.

3.2 Hyperparameter Selection

To explore the performance of the evolutionary algorithm for different parameter combinations, we performed a hyperparameter sweep over a range of values for the population size, crossover probability, mutation probability, and tournament size. For each parameter combination, we ran the algorithm 20 times to account for the stochastic nature of the evolutionary process.

3.3 Perturbation Model

To simulate catastrophes, we implemented a perturbation operator that could be called on a genome within a generation of the evolutionary algorithm. The perturbation operator receives three parameters, the number of perturbations, the size of perturbations, and the spatial distribution from which to sample the location of the perturbations. The function is called a number of times intra-generationally based on this first parameter. When called, a weighted random point is selected from within the bounds of the contiguous US where the weighting is the provided distribution. Despite this modular implementation, for the results in this paper, our runs used only a flipped Pareto distribution that weighted points to be selected from the eastern seaboard, as can be seen in figure **??**. After the point is selected, a radius corresponding to the size of perturbations parameter is calculated within which all facilities are removed from the genome. After these removals have taken place a number of times, the fitness of the genome is calculated and stored, and then then genome itself is restored to its state prior to the perturbations with the new fitness.

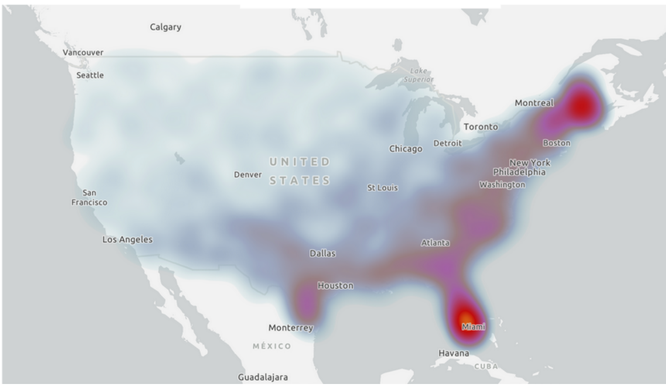

Figure 1: Flipped Pareto distribution overlaid on a contiguous map of the US from which to draw perturbation locations

4 RESULTS

4.1 Hyperparameter Sweep

In order to determine the best set of parameters we performed a hyperparameter sweep over the Cartesian product of the following hyper parameter values. The hyperparamter values swept over are listed in table.

Hyperparamter	Values
Number of elements to mutate	1,3,5
Crossover points	1,2,5,10,20
Num tournament winners	1,2,3,4,5
Tournament size	5,10,20

The best set of parameters is as follows: number of elements to mutate = 1, crossover points = 5, tournament size = 20, and number of tournament winners = 1. The results indicate that the evolutionary algorithm benefits from a high selection pressure, indicating that the landscape is smooth.

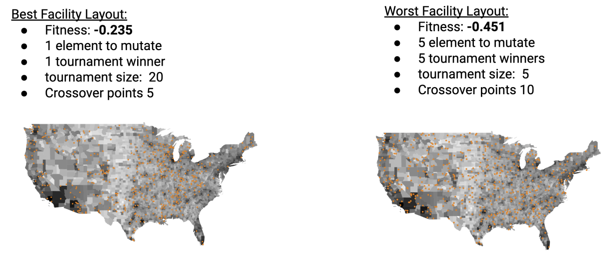

Figure 2: Resulting genomes from highest and lowest fitness runs of the parameter sweep. Facilities are represented as orange dots over a choropleth map of county population. The high fitness run resembles a real facility layout with facilities clustering around population centers while the low fitness map shows a uniform spread of facilities.

4.2 Scaling Relations

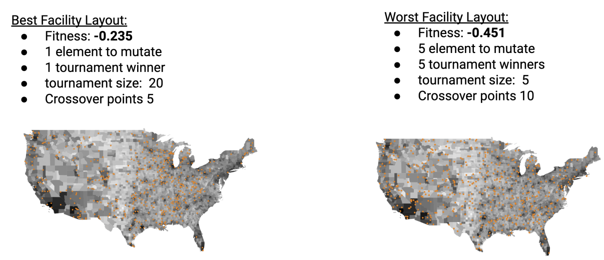

Figure 3: Resulting genomes from highest and lowest fitness runs of the parameter sweep. Facilities are represented as orange dots over a choropleth map of county population. The high fitness run resembles a real facility layout with facilities clustering around population centers while the low fitness map shows a uniform spread of facilities.

We compare the highest and lowest fitness individuals in the population and measure their scaling relationships between facility density and population density. As stated above, the optimal facility placement will result in a scaling relationship with an exponent of 2/3. The scaling relationships are shown in Fig **??**. We see that the best fitness individual scales with an exponent of 0.74 while the lowest fitness individual has a scaling relationship of 0.399. This indicates that our evolutionary algorithm performs well as a metaheuristic solution to the p-median problem which arrives at solutions close to the optimal one. Visual inspection of Fig. **??** shows the difference in the resulting genomes. The highest fitness solution clusters facilities around urban centers while the lowest fitness facility distributes facilities more or less randomly throughout the space.

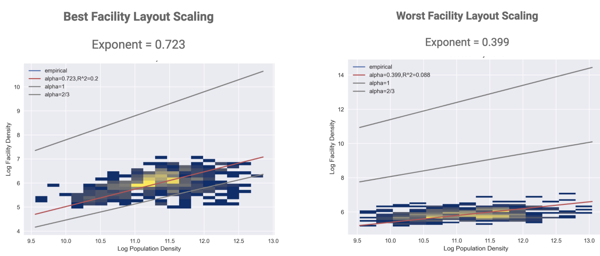

Figure 4: Resulting scaling relationships for highest and lowest fitness individuals. The lowest fitness individual shows little scaling with the population with a scaling exponent of 0.399. On the other hand, the highest-fitness individual has a scaling exponent of 0.74, close to the theoretical value of 0.66

4.3 Perturbation Analysis

We additionally run our model with perturbation and survey the results. We select the highest fitness genomes from runs of the algorithm with no perturbation as well as runs of the algorithm with perturbation. We use two tests to quantify the robustness of our solutions. First, we apply a series of perturbations to each genome and measure the number of facilities that have been removed. We also measure the drop in fitness resulting from the perturbations. A more robust facility layout will minimize the drop in fitness resulting from perturbations.

Fig **??** depicts a comparison of the average drop in fitness between a genome evolved with perturbation and a genome evolved without perturbation. Each genome is subjected to 5,000 independent perturbations of size 10 and the drop in fitness from each is measured.

The genome with perturbation had a higher mean than the genome with perturbation and appears to be bimodal. In this plot, both parameters use the hyperparameters shown for the best genome in Fig **??**. The genome evolved with perturbation was evolved with 20 perturbations of size 10. The genomes are evaluated by applying 5,000 perturbations of size 10 and measuring the drop in fitness from each. The mean drop in fitness due to perturbation on the genome evolved without perturbation is approximately 10% higher than the drop in fitness on the genome evolved with perturbation. A KS test rejects the null hypothesis that both samples are drawn from the same distribution with a p-value of 0.049. This indicates that the mean is in fact higher. Our results suggest that facilities evolved with perturbation do in fact exhibit robustness.

5 DISCUSSION

Our results are a promising start. We show that evolutionary algorithms are a viable alternative solution to the p-median problem which arrives at near-optimal solutions. We also show that the evolutionary algorithm is able to evolve facility placements that are robust to perturbations. Our results regarding perturbations are preliminary and a much larger and more comprehensive study is needed to properly quantify the true effect of perturbation and the parameters at which perturbation can successfully evolve robust solutions.

There are several extensions to this research we would like to undertake. The current analysis only examines the perturbations drawn from a Pareto distribution. We intend this analysis to examine the different methods of location selection for perturbations. This includes drawing perturbations from a uniform distribution as well as selecting facilities with high populations within their Voronoi cells preferentially. We also intend to compare our results against the current benchmark optimization method for the p-median problem - simulated annealing.

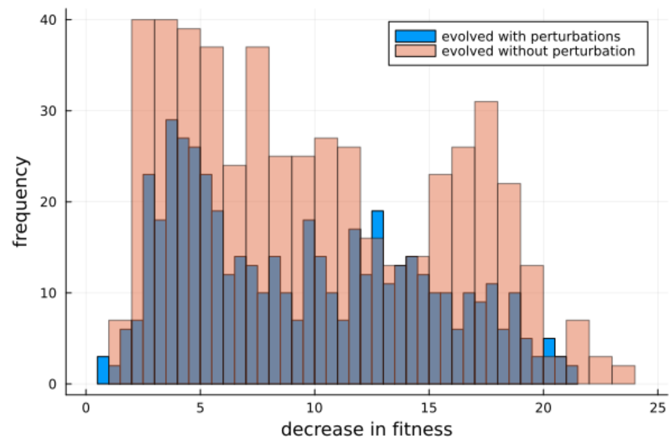

Figure 5: We apply 5,000 perturbations to each genome and measure the resulting drop in fitness from each perturbation. We plot the results for the genome evolved with perturbation and the genome evolved without perturbation.

6 CONCLUSION

In conclusion, the evolutionary algorithm we developed for the p-median problem was able to find near-optimal solutions. Our results reflected the $\frac{2}{3}$ scaling that we expected to see in an optimal solution to the problem. Using larger tournament sizes, a fixed mutation probability of 1, and a lower number of crossover points generally improved the performance of the algorithm. Our results also showed that the evolved solutions were able to trade off some initial fitness in order to have less change in fitness after a catastrophic perturbation. Overall, our study demonstrates the potential of using evolutionary algorithms to solve the p-median problem, and to evolve genomes while accounting for robustness of the evolved solutions to external shocks.

7 ACKNOWLEDGEMENTS

We appreciate the feedback and edits from Professors Nick Cheney, Peter Dodds, and Chris Danforth of the University of Vermont.

8 COMPETING INTERESTS

The authors of this paper have no competing interests.

REFERENCES

[1] Nusaybah Alghanmi et al. "A Survey of Location-Allocation of Points of Dispensing During Public Health Emergencies". en. In: *Frontiers in Public Health* 10 (Mar. 2022), p. 811858. ISSN: 2296-2565. DOI: 10.3389/fpubh.2022.811858. URL: https://www.frontiersin.org/articles/10.3389/fpubh.2022.811858/full (visited on 12/16/2022).

[2] J. M. Carlson and John Doyle. "Complexity and robustness". EN. In: *Proceedings of the National Academy of Sciences* 99.suppl_1 (Feb. 2002). Company: National Academy of Sciences Distributor: National Academy of Sciences Institution: National Academy of Sciences Label: National Academy of Sciences Publisher: Proceedings of the National Academy of Sciences, pp. 2538–2545. DOI: 10.1073/pnas.012582499. URL: https://www.pnas.org/doi/abs/10.1073/pnas.012582499 (visited on 10/24/2022).

[3] S. M. Gusein-zade. "Alternative explanations of the dependence of the density of centers on the density of population". eng. In: *Journal of Regional Science* 33.4 (Nov. 1993), pp. 547–558. ISSN: 0022-4146. DOI: 10.1111/j.1467-9787.1993.tb00848.x.

[4] Jaegon Um et al. "Scaling laws between population and facility densities". In: *Proceedings of the National Academy of Sciences* 106.34 (Aug. 2009). Publisher: Proceedings of the National Academy of Sciences, pp. 14236–14240. DOI: 10.1073/pnas.0901898106. URL: https://www.pnas.org/doi/10.1073/pnas.0901898106 (visited on 09/23/2022).

[5] Tong Zhou, J.M. Carlson, and John Doyle. "Evolutionary dynamics and highly optimized tolerance". en. In: *Journal of Theoretical Biology* 236.4 (Oct. 2005), pp. 438–447. ISSN: 00225193. DOI: 10.1016/j.jtbi.2005.03.023. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022519305001323 (visited on 10/13/2022).