

YZV 202E Optimization for Data Science Project

Proposal Efficient Selection of Green Space Locations in Istanbul Project Report

Yusuf Solmaz

Computer and Informatics Department
Istanbul Technical University
solmaz22@itu.edu.tr
150220306

Necip Fazıl Doğan

Computer and Informatics Department
Istanbul Technical University
doganne23@itu.edu.tr
150230311

İbrahim Bancar

Computer and Informatics Department
Istanbul Technical University
bancar22@itu.edu.tr
150220313

Abstract—This report analysis Efficient Selection of Green Space Locations in Istanbul project. The project aims spread green space over Istanbul using Genetic Algorithm and Particle Swarm Optimization. Current green space size, population, air quality and transportation accessibility conditions considered in county determination process. The objective function aims to improve the state of life with balance of these factors. The objective function has also cost constraints. The report analysis convergence success of PSO and GA. Also, the report examines reaction of objective function under different weights.

I. PROBLEM DESCRIPTION

Green spaces are very important for quality of life, sustainability and public health. In metropolis like Istanbul, the population increases and urban sprawl causes green spaces to spread inefficiently. As a result, air pollution increases, heat distribution affects negatively and like this problem. These problems have negative effects on public health.

The project aims determine new green space locations and sizes based on population, air quality, current green spaces and transportation accessibility. Main problem is determination of county need green spaces and how much size under cost constraints.

II. METHOD FORMULATION

A. Data Collection and Process

- **Current Green Space**
Data including current green space information as geojson file gathered from official IBB website. This data processed for split data with respect to counties and total green space. Final data includes total green space area for each county.
- **Population**
Population data is collected from IBB website and categorized each counties population
- **Air Quality**
Air quality data is collected from IBB api. AQI scores are saved hourly for all days between 2024-2025 from 28 air quality measurement center. Each center takes average of

the yearly AQI value. Rest of counties take value from the closest center.

- **Public Transportation**
Number of railway, minibuss and taxi stations are gathered from IBB dataset. These stations are weighed in support of railway more to determine transportation score for all counties.

B. Objective Function

$$\text{Minimize } f(x) = \sum_i^n \frac{(W_1 \cdot AQI_{i,norm} + W_2 \cdot T_{i,norm})P_i}{GS_i + x_i + \epsilon}$$

AQI: Air Quality Index

T: Transportation Score

P: Population

GS: Current Green Space

x: New Green Space

AQI and T values are used under linear normalization. ϵ value is for avoiding divided by zero error.

Objective function give sum of the all counties performance value. This value has negative correlation with new and current green space size. Also, this value increase with respect to air quality, transportation and population scores. To minimize this function, the x values become larger with respect to the other parameters. W_1 and W_2 weights determine which parameter is more important between air quality and transportation. Penalty is used for county exceed the constraints.

C. Constraints

- **Total Green Space Limit**

$$\sum_{i=0}^n x_i \leq 3,000,000$$

This constrain is for prevent exceeding cost limit with too much rising of green space

- **County Green Space Limit**

$$x_i \leq 0.5 \cdot GS_i$$

- Positive Space

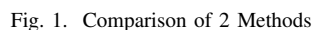
This constraint is for prevent negative space for right result.

1) *Genetic Algorithm*: New green space vector x is used chromosome in GA.

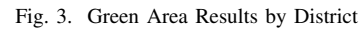
- ### III. REAL-WORLD APPLICATION

IV. EXPERIMENTAL EVALUATION

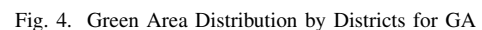
A. Comparison of 2 Methods



The graphs below show the amount of green space per capita before adding green space and how much green space the PSO and GA provide for each district per capita.



When we examine how the two algorithms distribute green areas to the districts, we see that GA distributes green areas to the entire map, while PSO distributes all green areas to a few districts and does not give any green areas to other districts. The green area distributions of GA and PSO are shown on the map in the images below.



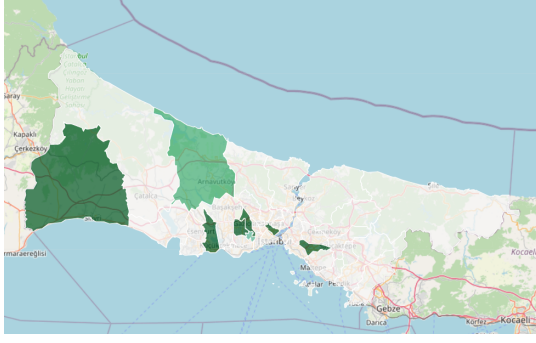


Fig. 5. Green Area Distribution by Districts for PSO

V. EFFECT OF W1 AND W2 VALUES ON DISTRIBUTIONS

When the objective function is examined, it is expected that the W1 value will provide more green areas to districts where air quality is poor and the W2 value will provide more green areas to districts where transportation is easy. However, when the W1 and W2 values are increased or decreased, instead of this expected result, both algorithms behave more randomly. Despite this, it is seen that the behavior of concentrating on certain areas is more frequent in scenarios where one of the W1 and W2 values is high and the other is low, but this does not create a clear pattern.

In addition, it is seen that the increase in the W2 value generally causes the best scores to increase, which is the aim of the algorithms to minimize this.

VI. CONCLUSION

In this study, the issue of fair and effective allocation of green spaces across the districts of Istanbul was addressed by incorporating multiple real-world factors, including population, existing green areas, air quality, and accessibility to public transportation.

Two heuristic optimization methods, Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), were applied and compared under the same circumstances in order to address this issue. While GA gives slightly better results than PSO, it is seen that PSO converges much faster. When we look at the behavior of the algorithms, we see that GA provides green areas to a wider region, while PSO concentrates on some areas.

Additionally, it was examined how sensitive the objective function was to the weighting parameters W1 and W2, which stand for the respective significance of accessibility to transportation and air quality. It was discovered that imbalances in the weight arrangement frequently resulted in more polarized allocation outcomes, even though totally predictable behavior was not always seen.

This study demonstrates how heuristic optimization techniques can be applied to solve challenging urban planning issues while taking practical limitations into account. The suggested approach offers a data-driven framework that could help policymakers support sustainable development and environmental fairness in urban areas. Future work may focus

on the integration of temporal planning models, stakeholder preferences, or real-time data to further enhance the model's applicability.