2024-05-01

Using Genetic Algorithm And Geographic Information System For Equity-Driven Transportation Planning And Analysis

Emiliano Del Rio Reyes
University of Texas at El Paso

USING GENETIC ALGORITHM AND GEOGRAPHIC INFORMATION SYSTEM FOR EQUITY-DRIVEN TRANSPORTATION PLANNING AND ANALYSIS

EMILIANO DEL RIO REYES

Master's Program in Civil Engineering

APPROVED:

Adeeba A. Raheem, Ph.D., Chair

Ruey L. Cheu, Ph.D.

Thomas Horak, Ph.D.

Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School

USING GENETIC ALGORITHM AND GEOGRAPHIC INFORMATION SYSTEMS FOR EQUITY-DRIVEN TRANSPORTATION PLANNING AND ANALYSIS

by

EMILIANO DEL RIO REYES

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Civil Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

May 2024

Abstract

Efforts to address equitable access in public transit planning have gained momentum, spurred by incentives for Metropolitan Planning Organizations (MPOs) and State Departments of Transportation (DOTs). However, traditional strategies often fall short of meeting the needs of disadvantaged communities, particularly in underserved areas. This study presents a pioneering methodology leveraging Genetic Algorithms (GA) and Geographic Information Systems (GIS) to optimize bus stop placement, aiming to enhance equitable access in public transit systems. Focusing on Route 16 of Sun Metro in El Paso, Texas—a critical feeder route linking the Westside Transit Center to Upper Valley neighborhoods—the research commences with a comprehensive analysis of demographic data and existing transit conditions to pinpoint disparities and accessibility challenges. By harnessing GA and GIS, the study proposes solutions tailored to equity factors, resulting in notable improvements in accessibility metrics. The research underscores the imperative of modernizing evaluation methodologies and integrating emerging technologies. Despite encountering challenges such as data availability constraints, computational demands, and the dynamic nature of urban environments, the study advocates for developing adaptive models. This research contributes significantly to advancing equitable transit systems and practices, offering valuable insights and a replicable methodology to enhance accessibility and equity in public transportation networks.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Traditional methods of transit planning often overlook crucial demographic factors (Linovski et al., 2022), resulting in disparities in access to public transportation services. This accessibility gap can exacerbate the marginalization of populations already experiencing disadvantageous conditions. The efficiency of transit systems should not be evaluated only on time, cost, or connectivity but also based on the needs of those who use it and those who need it. Addressing the disparities faced by disadvantaged groups requires a comprehensive understanding of their needs and preferences, offering a potential avenue to alleviate some of the socio-economic burdens they endure. It is crucial to look towards a fair distribution of transportation resources, which provides more options to those who lack travel choices and possibly reduces their travel time (El-Geneidy et al., 2016).

Over the years, the role of public transit has evolved beyond mere transportation infrastructure, morphing into a crucial cornerstone of societal welfare. Across most cities, barring some exceptions, commuters reliant on public mass transit heavily depend on its availability to access essential services and goods (Garrett & Taylor, 2012). Unfortunately, there has been an observed inverse correlation between areas predominantly inhabited by low-income households and minorities, as well as the accessibility of transit services. In these marginalized neighborhoods, where economic hardship and social marginalization often converge, access to reliable public transportation emerges as a pivotal lifeline. Yet, these communities bear the brunt of inadequate transit infrastructure, grappling with limited service coverage, prolonged wait times, and substandard connectivity. This observed disparity underscores a fundamental challenge in contemporary urban planning: the persistent exclusion of marginalized populations from the benefits of accessible and efficient transit networks. The ramifications of this inequity ripple far beyond the realm of transportation, permeating various facets of daily life, from access to employment and education to healthcare and social opportunities (Casas, 2007; Preston and Rajé, 2007; Lucas, 2012; Pereira et al., 2017).

This study proposes integrating socio-economic factors and technologies to enhance public transit planning. Drawing insights from Genetic Algorithms (GA) and Geographic Information Systems (GIS) analysis, the research aims to achieve the following objectives:

1) Analyze transportation accessibility in the County of El Paso, TX, based on various socio-economic factors.
2) Develop a hybrid model that combines Genetic Algorithms and GIS analysis to select optimal bus stop locations that align with the needs of the local demographics.
3) Analyze the potential impact of the proposed optimization on the local ridership volume.

In essence, this research aims to foster more equitable and accessible transit systems. By incorporating new technologies and repurposing tools, this thesis seeks to include equity as a fundamental indicator of transit efficiency. By prioritizing the needs of underserved communities and leveraging data-driven insights to inform decision-making processes, the proposed hybrid model holds the potential to catalyze positive transformative changes within public transportation systems, ultimately leading to more equitable outcomes for all stakeholders involved.

The thesis document consists of 5 chapters and an appendix. Chapter 1 introduces the research problem and objectives and outlines the chapters in this thesis. Chapter 2 is focused on providing a comprehensive review of the extant literature concerning transportation equity, particularly on underserved communities and the evolution of transportation planning tools. Chapter 3 delves into the research methodology, followed by data analysis in Chapter 4. Finally, Chapter 5 provides concluding remarks and directions for future research. Appendix A provides the ArcPy code for the proposed model to perform geographic data analysis.

Chapter 2: Literature Review

Transportation equity is a topic that has been gaining attention in recent years. This heightened focus can be attributed, in part, to initiatives such as the Justice40 Initiative, which have galvanized efforts to prioritize the needs of disadvantaged communities within transportation planning and policy (Walls et al., 2024). Similar programs like the U.S. Department of Transportation's (DOT) Grant Programs RAISE and Reconnecting Communities signal a tangible commitment to address the historical inequities and foster more inclusive transportation systems. The imperative to address transportation equity stems from recognizing that access to reliable and efficient transportation services is not merely a matter of convenience but a fundamental determinant of individual and community well-being. The ability to secure gainful employment, access essential healthcare services, and attend educational institutions hinges on one's ability to navigate the transportation network easily and affordably (Pereira & Karner, 2021). In essence, equitable transportation provision serves as a linchpin for socioeconomic mobility and opportunity, amplifying the voices of marginalized populations and narrowing the pervasive gaps in access and opportunity.

DEFINING TRANSPORTATION ACCESSIBILITY AND EQUITY

Transportation Accessibility

The concept of accessibility in transportation is continually evolving to reflect the shifting needs and circumstances of contemporary society. While definitions may vary, there is a general consensus that accessibility pertains to an individual's ability to reach goods, services, and rights, serving as a pivotal criterion for evaluating urban transportation systems. Indeed, accessibility plays a crucial role in advancing social justice within urban environments (Jamei et al., 2022). Jomehpour and Smith-Colin (2020) have underscored the historical challenge of grappling with multiple terms for accessibility, leading to the adoption of redundant criteria or outdated factors in transportation planning and policymaking processes. Malekzadeh & Chung (2020) have two distinct dimensions of accessibility: "active accessibility," which pertains to an individual's desire and ability to reach a specific destination or engage in a particular activity, and "passive accessibility," referring to the capacity of multiple users to access a given location. These contrasting dimensions underscore the multifaceted nature of accessibility and its implications within this domain. To comprehend the intricacies of accessibility, it is imperative to identify the factors that influence it. Although transportation demand and activity stand out as prominent determinants, as highlighted by Litman (2008), demographics, purpose, destination, time, mode, and distance are among the key factors considered in assessing accessibility within transportation systems (Litman, 2008).

X. Chen (2018) expands the concept of accessibility by classifying "Transit Accessibility" into different categories based on contextual factors or objectives. These categories include destinations' activity types (e.g., jobs, school, shopping), spatial dimensions (e.g., local vs network accessibility), temporal dimensions (e.g., peak hours), and components (e.g., land use, transportation infrastructure, individual characteristics). This systematic classification highlights how varying perspectives and considerations alter the definition and treatment of accessibility.

Transportation Equity

Equity in the transportation context gained momentum in the United States, spurred by historical events and policies that have shaped discussions around social justice and civil rights. Iconic examples include the 1955 Montgomery Bus Boycott, which served as a catalyst for the Civil Rights Movement (Inwood et al., 2015). Furthermore, policies such as the enactment of Title VI of the Civil Rights Act of 1964 marked early attempts to address inequities in transportation accessibility. Recently, initiatives like Justice40 (2021) have reignited interest in addressing transportation equity as an ongoing challenge (Antipova et al., 2020). The U.S. Department of Transportation (USDOT) has identified clean transit as one of the investments covered by the initiative. It has three main components for its implementation: Understanding a community's needs through public engagement, Understanding how a community is affected by a lack of transportation options and investments, and Understanding the benefits a project may create. Initiatives like this at the Federal level underscore the importance of ensuring equitable access to transportation resources and opportunities, reflecting a continued commitment to advancing social justice within the transportation sector.

The evaluation of transportation systems performance also has evolved over time. As in previous decades, the primary focus of the evaluation was centered on enhancing speed, minimizing travel time, and alleviating congestion (Litman, 2022; Manaugh et al., 2015). However, the emergence of Environmental Justice principles has precipitated a notable shift in how the performance of these systems is measured. These initiatives emphasize the need to evaluate transportation systems through a lens that considers their impact on social equity, environmental sustainability, and community resilience.

However, there is no clear consensus on defining or quantifying equity, suggesting that we are likely in the nascent stages of integrating transportation equity into urban planning and policymaking processes. Bruzzone et al. (2023) discuss two definitions utilized over the years. The first definition states equity as "the morally proper distribution of benefits and costs (burdens) among members of society." In this perspective, the focus shifts towards a morally-driven interpretation of equity, considering the societal members to whom these benefits and burdens should be allocated and assessing the moral acceptability of such distribution. The other definition states that equity focuses on " providing a wider variety of choices to people who have fewer ones." This definition is particularly relevant in the context of transportation planning, as it emphasizes the need for policymakers to expand service accessibility, especially in less densely populated areas, in order to offer a greater range of transportation options to those with limited choices. The Justice40 initiative, established by the United States Biden administration, follows the path of the second definition discussed previously. Since it focuses on distributing at least 40 percent of overall benefits to disadvantaged communities. This includes, but is not limited to, infrastructure, climate change investment, clean transit, remediation, and clean water (Sotolongo, 2023). This initiative tries to distribute society's benefits more equitably by giving priority to those individuals and groups that have historically been underserved.

Litman (2018) introduces the concept of Horizontal and Vertical Equity by providing a different framework for understanding and assessing equity in transportation planning and policymaking. Horizontal equity focuses on the equal distribution between individuals or groups with the same characteristics. This implies that everyone should be treated equally and that no one should be favored. Vertical equity, which he also refers to as *Social Justice, Environmental Justice, and Social Inclusion*, focuses on addressing disparities in impact among individuals or groups with differing characteristics. To address these disparities and promote greater equity, Vertical equity advocates for targeted interventions to prioritize the needs of disadvantaged communities. This may involve implementing various measures, such as affordability enhancements to make transportation more accessible to low-income individuals, improving transit service in underserved areas, and providing support services to facilitate mobility for vulnerable populations, such as seniors and people with disabilities.

UNDERSERVED COMMUNITIES

Understanding transportation equity necessitates central consideration of disadvantaged communities or underserved populations, given that transportation is a fundamental public service and necessity. Identifying and prioritizing these communities constitutes a foundational aspect of equitable transportation planning, ensuring the fair distribution of benefits and resources. The California Public Utilities Commission defines disadvantaged communities as "areas throughout California which most suffer from a combination of economic, health, and environmental burdens." These areas encompass factors such as poverty, elevated unemployment rates, and pollution, indicating a disproportionate burden on residents' quality of life. Recognizing the significance of this definition underscores the imperative to identify these communities and discern their primary needs.

Ong et al. (2021) studied the differences among disadvantaged neighborhoods of Los Angeles and San Joaquin. They compared the demographics between disadvantaged and non-disadvantaged neighborhoods in both zones and found similar results. Hispanic and Black populations predominated in disadvantaged neighborhoods, whereas non-Hispanic White populations were more prevalent in non-disadvantaged neighborhoods.

Most of the time, the areas where disadvantaged populations tend to reside have low-quality mobility infrastructure, such as sidewalks or cycling lanes, which results in them snowballing into even lower conditions of education, poverty, or health (Nicoletti et al., 2023; Pereira & Karner, 2021). Similar results were shown in another study conducted by Ward & Walsh (2023) to understand the effects in transit-disadvantaged individuals. Through interviews, they found out that unreliable access to transportation systems limits the options of these underserved communities since these individuals often make sacrifices to meet their bare minimum needs, with repercussions extending beyond adults to affect children as well. Examining disadvantaged communities and underserved populations shows the challenges faced by these groups, and how the lack of accessible transportation services prevents them from achieving an acceptable quality of life. These communities bear a disproportionate economic, health, and environmental burden, further compounded by deficient transportation infrastructure. The findings highlight the need to prioritize these communities by identifying and understanding their need for enhanced transportation equity.

Transit Planning

The existing literature suggests that there are three main components of planning for transit systems: time/speed, cost-benefit, and accessibility (Adli & Chowdhury, 2021; Liu & Cheng, 2020; Murray & Wu, 2003). It is important to understand that while we may analyze these variables separately, they are not totally independent variables. Cost-benefit analysis is a common approach for determining bus stop locations as it assesses the balance between transportation costs and the benefits derived from it, with benefits defined as accessible opportunities or activities. These benefits are influenced by various factors, including distance and bus stop usage. Adli & Chowdhury (2021) highlight that performing cost-benefit analysis (CBA) is the usual practice as it provides insights into whether the benefits of a particular option outweigh its associated costs. Due to its widespread use for decision-making, it underestimates potential social impacts. For instance, reducing car ownership can substantially decrease the demand for parking spaces and mitigate environmental impacts. That is why they propose using social CBA or multi-criteria analysis to meet the social aspects and have a broader understanding of accessibility.

The Federal Transit Administration recommends (2011) an ideal 400-meter walking distance to the bus stops. However, the feedback from transportation planners and community members suggests that the pedestrian catchment area should not be based on mileage but should be analyzed on a case-by-case basis. Murray & Wu (2003) mention that the distance between bus stops changes based on the location. FTA suggests a typical spacing of 230 meters, but their research shows that most urban transit agencies typically have distances ranging from 200 to 600 meters. Currently, agencies' suggestions are not followed in their entirety and there may be divergent opinions or objectives between urban planners and transit agencies, which suggests that there is room for improvement in how transit accessibility and bus stop performance are measured and analyzed.

Additional feedback to the FTA suggests that their recommended 15-minute walking time to a bus stop should be extended by 30 minutes, as users may be willing to walk for longer periods, potentially reducing emissions. However, this contradicts findings by Murray & Wu (2003) who argue that a 5-minute walking time is reasonable. Speed plays a pivotal role in determining bus stop locations, particularly in bus rapid transit (BRT) systems, where walking distance and spacing may be sacrificed to enhance speeds and minimize waiting times.

Concepts like Transit Oriented Development (TOD) prioritize walkability and the transformation of car-centric neighborhoods. Ibraeva et al. (2020) introduce the notion of maintaining an average walking distance of 600 meters from bus stops to points of interest within this framework. There can be secondary areas with a maximum distance of 1.6 kilometers. In general, TOD considers increasing the conditions of the users' surroundings to promote active transportation. The key behind this concept is having well-developed public transportation, which is convenient for users, resulting in the reduction of traffic congestion by diminishing the reliance on personal vehicles (Sharma et al., 2024).

Linovski et al. (2022) compared three major cities and analyzes the construction of BRT systems. They use buffer zones to assess accessibility to these systems for each census tract. By utilizing demographic data, they identify high-need zones to measure the accessibility of these areas. Their findings indicate that, in one instance, incorporating high-need census tracts could have significantly enhanced accessibility. This is attributed to including routes based on service to existing land use and projected ridership rather than solely on actual ridership. Murray & Wu (2003) propose meeting the needs of the community and maximizing ridership. It is suggested that higher speeds can increase utilization, while stop spacing may not be as important as efficient routes and placements.

An essential aspect of equitable transportation planning involves identifying and involving stakeholders, particularly through community engagement. As previously highlighted, recognizing and comprehending the unique needs of each community is crucial, as there is no one-size-fits-all solution. This emphasis on community engagement is paramount, given that the primary goal of transit-related projects should be to enhance the quality of life for the public (Erkul et al., 2020). Therefore, the transit planning process should consider incorporating the values and perspectives of communities, meeting their needs and expectations while also addressing their concerns.

MEASURING EQUITY AND ACCESSIBILITY

Planning transportation systems is a multifaceted process that constantly requires careful consideration of various factors and methodologies. Wei et al. (2017) proposed a method to evaluate operational efficiency and access equity that implements a combination of Geographic Information System and multi-objective spatial optimization techniques. This model uses variables like ridership and investment labor for each route to determine operational efficiency, while equity is determined by the service coverage for disadvantaged populations. Their analysis indicates that adopting these tools enables us to examine equity and its trade-offs with operational efficiency. Despite being constrained by available data; this method remains robust for assessing accessibility and equity.

Other GIS-based analysis approaches take into consideration the network itself and the information related to its operation (Kaplan et al., 2014). The main analysis components are transit connectivity, location-based and potential-accessibility measures, and equity assessment. The underlying concept of this model is to ascertain the network's connectivity, encompassing origins to multiple destinations and multiple origins to a single destination. Additionally, the calculation of Gini coefficients is incorporated to gauge the extent of inequality. Another method involves the concept of transit deserts, which aims to identify areas for enhancing the current transit system by assessing the disparities between demand and supply. It begins by identifying the transit-dependent population or demand, typically based on demographics for disadvantaged populations and individuals with limited mobility. The next step involves estimating transit supply, achieved through the calculation of a comprehensive public transit accessibility score using various factors and indicators (Jiao & Dillivan, 2013; Jomehpour Chahar Aman & Smith-Colin, 2020b).

Finio et al. (2024) provide an understanding of mapping equity and its current state as more MPOs and state agencies develop visualization tools to represent the gathered data. Creating the equity and opportunity maps requires gathering socio-economic, demographic, and public health indicators from sources such as the U.S. Census and state/local agencies. The authors highlight indicator fatigue as a common challenge, wherein the abundance of data can complicate the process of prioritizing variables and necessitate additional effort to update information. This research highlights the effectiveness of these maps to show disparities, leading to an informed decision-making process. However, there is a clear need for simpler, more straightforward maps, as the most useful maps are often the simplest ones.

Community engagement is a key component in developing these datasets, as community members and residents often contribute to data collection and validation (Finio et al., 2024). The authors of this study highlight Portland's inclusive map-building process, which involves various stakeholders determining the indicator list. However, one of the challenges faced was the divergence in priorities and objectives among stakeholders, leading to some members feeling disheartened as their priorities were not reflected in the outcome.

GEOGRAPHIC INFORMATION SYSTEM (GIS)

The Geographic Information System (GIS) has become an indispensable tool for transportation planning and equity analysis. The tool interprets data through visual representations and conducts analyses essential for decision-making. GIS facilitates the integration of various datasets related to transportation infrastructure, demographic information, socioeconomic factors, and environmental impacts (Droj et al., 2021; Zannat et al., 2020). Effective use of GIS can greatly assist planners and policymakers in mapping collected data from users and transportation networks. These maps can later be used to identify underserved areas and assess the current or possible resource distribution. GIS enables the exploration of potential equity impacts by allowing an informed decision-making process that prioritizes disadvantaged communities.

The GIS platform emerged around fifty years ago primarily to automate map production, initially lacking analytical capabilities, as its primary objective was visual data presentation. However, today, it has diverse analytical functionalities, leveraging both spatial and non-spatial data. For instance, it can utilize historical data to discern patterns and spotlight problematic zones at particular times. Urban planning and transportation experts are now seamlessly integrating spatial analysis into conventional methodologies (Droj et al., 2021)

Yona et al. (2021) implemented a GIS-based analysis to understand and propose solutions to user complaints in transportation systems. Their results suggest that planners and authorities can benefit from the proposed model as mapping can help them understand the geographical distribution, changes over time, and the impact of service improvements. Implementing the suggested solutions can significantly enhance authorities' ability to gather feedback from passengers and, through advanced data analytics, gain insights into the system.

The study performed by Zannat et al. (2020) aims to understand the relationship between environmental conditions and accessibility of public transportation. Using a GIS-based analysis approach incorporating Frequency Ratio (FR) and Analytical Hierarchy Process (AHP), researchers analyzed eight environmental factors, including slope, elevation, land use, and public transport service area. Study findings could potentially help planners monitor existing facilities and identify areas needing improvement to enhance accessibility. Conducting spatial analysis provides insights into the system's performance according to the selected parameters. For instance, through place-based measurements, analysts can evaluate the actual accessibility of a specific bus route (Higgins et al., 2022) .

GIS can also be used to evaluate potential accessibility, which refers to the capacity of the users to reach their destinations using the available transportation methods. It measures how well or efficiently a transportation system serves its users' needs in terms of origin-destination (e.g., home-school, home-work, work-recreational facilities). It is common to see this type of analysis when planning for the construction of new services, showing the capacity of customers to use the services. Kompil et al. (2019) analyzed the potential accessibility of what they call generic services using a script in ArcGIS, MATHLAB, and GeoDMS. Their objective was to determine the extent to which the population has access to these services. In a similar study, Goliszek (2021) examined the accessibility of both private and public transportation, comparing their potential accessibility throughout the day and during rush hours. The model utilizes publicly available data from the Internet, including General Transit Feed Specifications (GTFS) for transit information and the Google Maps API for private vehicle data. The model used potential-gravity methods to calculate accessibility to the users' destination based on time.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN TRANSPORTATION

The use of Artificial Intelligence (AI) and Machine Learning (ML) technologies has been increasing in transportation planning, facilitating decision-making processes by addressing complex challenges. Leveraging AI and ML capabilities enables transportation planners, authorities, and researchers to gain deeper insights from gathered data, uncovering novel patterns and disparities. This, in turn, facilitates the development of precise, targeted solutions. Genetic Algorithms (GA) is a subfield of AI that uses evolutionary concepts to learn and create solutions based on criteria and specifications. Among many applications, problem optimization stands out prominently. S. Chen et al. (2023) used GIS and GA to analyze problems within Toronto's public transit system to optimize the placement of bus stops, taking into consideration the distance between stops, cost, and other pedestrian factors. The results are compared to two other algorithms, Simulated annealing and PSO. While they found that GA performed much better, the drawback is the time required to reach the final solution, which is necessary for numerous generations.

GAs have also been used in vehicle routing problems. Abbasi et al. (2020) proposed implementing the GA model to solve the Travel Salesman Problem (TSP). Their model is based on existing GA solutions for this problem. In this study, the authors analyzed the possibility of utilizing a Parallel Genetic Algorithm (PGA), which has the characteristic of processing multiple solutions simultaneously. This is achieved by dividing the workload (number of solutions) between multiple cores. Similarly, PGAs performance in Traffic Problems has been tested to provide an adaptive signal dynamic speed control, which controls the time depending on each situation (Abu-Lebdeh et al., 2016). It considers the number of arterial links, queue length, volume of arriving traffic, and link speed. Mesbah et al. (2011) proposed to use GA to relocate road space between private vehicles and transit modes. This approach allows the use in real-world scenarios due to its reasonable computational time, making it a potential decision-making tool to optimize transit priority at the network level.

GA-based solutions have also been applied to calculate effective routes in green transportation operations (Lin et al., 2014). The objective was to calculate travel distances and propose solutions based on the cost effectiveness of the proposed routes. Some of the parameters it considers are number of nodes, distance between nodes, total distance, and vehicle capacity.

Chapter 3: Methodology

Problem Context and Objectives

Transportation planning and policymaking increasingly consider equity indicators in their analyses, yet the direction remains uncertain. Existing methodologies lack clarity in measuring equity, and initiatives like Justice40, while raising awareness, have yet to offer models for assessing and alleviating inequalities. This study proposed a framework for evaluating bus stop locations with equity as the primary indicator. The existing literature (Adli & Chowdhury, 2021; Wei et al., 2017) suggest two main approaches to evaluating transit systems: macro-level analysis focusing on network connectivity and micro-level analysis examining individual routes or bus stops. Micro-level analyses tend to prioritize cost, accessibility, and passenger capacity, neglecting equity considerations. Conversely, macro-level analyses often identify zones accessible to communities but overlook actual transit system accessibility. The proposed model seeks to bridge this gap by providing a micro-level analysis of equitable accessibility and offering solutions to enhance access for underserved communities.

Incorporating new technologies can be advantageous as they offer novel solutions from perspectives overlooked by outdated methodologies. These models harness the adaptability of genetic algorithms to generate solutions and the robust capabilities of GIS in visualizing and analyzing spatial and non-spatial data. While these technologies have been utilized in previous studies, their integration with equity as the primary indicator or variable in this model promises a deeper comprehension of equitable accessibility to bus routes.

Area of Study

This research is focused on El Paso County, Texas, situated at the western edge of the state along the Rio Grande. Its strategic position as a major border crossing between the United States and Mexico results in a high volume of commuters traveling from Ciudad Juarez for work or study. As per the U.S. Census, El Paso City has a population of around 680,000 residents, with the county housing approximately 870,000 residents. It ranks as the sixth largest city in Texas and the 22nd largest in the United States. El Paso is part of the Paso del Norte metropolitan area, also recognized as the El Paso-Juarez-Las Cruces area. The area is distinguished by its unique cultural amalgamation, interdependence, and historical significance, making El Paso a vital center for commerce as a major trading hub and culturally for its continuous exchange with Ciudad Juarez.

PUBLIC TRANSPORTATION OPERATOR

Sun Metro, the public transit authority in El Paso, traces its roots back to 1881 when it began as a trolley service connecting El Paso and Juarez. Formerly known as Sun City Area Transit (SCAT), it now has a comprehensive network comprising 53 bus routes and one streetcar line, depicted in Figure 1. Additionally, Sun Metro manages 8 transit centers, serving as hubs for 49 bus routes. The system organizes its services into six distinct categories (Table 1):

BRIO is the Sun Metro's BRT service. It currently operates four corridors: Mesa, Alameda, Dyer, and Montana. During peak hours, it runs every 10 minutes and every 15 minutes during non-peak hours. Local routes are the normal bus routes running mainly on arterial streets with frequent stops; they provide wide coverage. It currently operates 31 routes that run every 30 to 120 minutes.

Circulators are bus routes with looped alignments to provide service to one or more defined areas, helping with first and last-mile service. Besides its shorter route, they tend to share almost all characteristics of normal routes. The main areas where they operate are Downtown and Cielo Vista/Gateway Blvd.

Feeder routes connect areas in the outermost areas of El Paso to transit centers, where they can transfer to other Sun Metro services. Like circulators, they are helpful for first—and last-mile connections. Sun Metro currently operates 12 feeder routes.

Express routes are intended to connect transit centers over long distances. Besides Route 59, they usually have a service that runs every 50 to 70 minutes with limited stops.

Streetcar is a 4.8-mile fixed circulator route with two main loops, Downtown and Uptown. It provides service to two transit centers, although its service is subject to limitations.

Table 1. Sun Metro Bus Routes by Service Type

Sun Metro Routes	
Type	***Routes***
Brio	205, 206, 207
Local	2, 7, 10, 14, 15, 24, 25, 32, 33, 34, 35, 36, 37, 50, 51, 52, 53, 54, 58, 61, 62, 63, 64, 65, 66, 67, 68, 69, 72, 74, 86
Circulator	4, 8, 21
Feeder	11, 12, 13, 16, 19, 43, 44, 46, 56, 60, 84, 89
Express	5, 6, 26, 59
Streetcar	500

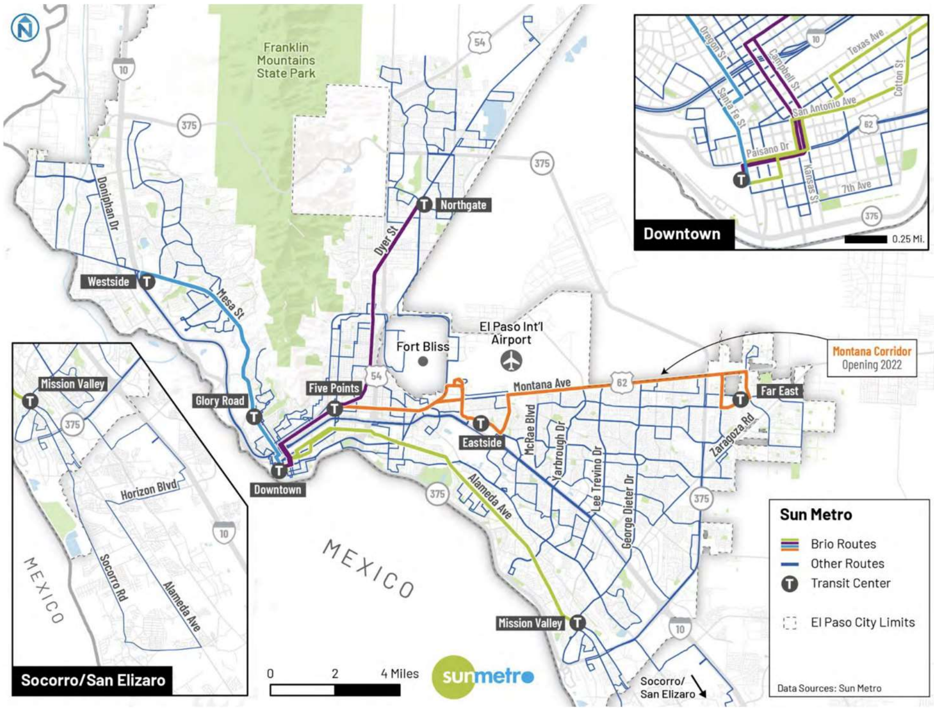

Figure 1. Sun Metro Bus Routes Layout (Source: State of System Report)

DATA AND RESOURCES

Three main datasets were collected to run the proposed model: El Paso census data, El Paso TIGER/Line Shapefile, and Sun Metro GTFS (see Table 2). These datasets are publicly available and can easily be downloaded.

Table 2. Datasets and File Type

Datasets	Format
Census Data (2020)	.csv
TIGER/Line Shapefile (2020)	shapefile
GTFS	Text

El Paso census data from 2020 serves as a foundation for this model, offering crucial demographic insights for two primary purposes. Firstly, it provides the geospatial reference (GEOID) required to understand the demographic composition and distribution within El Paso. Secondly, it provides location-specific indicators that are essential to assessing bus stop locations.

The data was retrieved from the Census Bureau website, specifically from the American Community Survey (ACS) and Decennial Census (DC). These datasets are available in multiple formats (geographic units), facilitating their integration into GIS for spatial analysis. For the proposed model, the necessary units are Census Block (CB) and Census Block Groups (CBG), as outlined in Table 3. The demographic indicators utilized are Total Population, Hispanic, White, Black, Asian, Poverty, Civilian Labor Force, and Civilian Labor Force Unemployed.

Table 3. Demographics, Origin, and Format

Survey	Data	Description	Geographic units
DC	P1_001N	Total Population	CB
DC	P1_003N	White	CB
DC	P1_004N	Black/African American	CB
DC	P1_006N	Asian	CB
DC	P2_002N	Hispanic	CB
ACS	B17021_002E	Poverty	CBG
ACS	B23025_003E	Civilian Labor Force	CBG
ACS	B23025_005E	Unemployed Civilian Labor Force	CBG

The Topologically Integrated Geographic Encoding and Referencing files provided by the U.S. Census Bureau serve as the visualization file of the census datasets. They do not contain demographic data, but the geographic entity codes (GEOIDs) are stored within them and can be used to reference the survey data. These maps are available at the national, state, or county levels. Usually, the files are divided by Census Blocks. However, other configurations, such as Census Block Groups or Census Tracts, are obtainable using the GEOIDs to dissolve the maps into the desired configuration.

The transit system data is available in the General Transit Feed Specification file, a dataset that contains static information about the system. This file format enables agencies to disseminate data to users in a manner that software tools can interpret and utilize. From this file, the coordinates for bus stops are retrieved, and shapefiles are outlined for bus routes.

GENETIC ALGORITHM IMPLEMENTATION

Genetic Algorithms (GA) are based on the process of natural selection, which falls under the Evolutionary Algorithms (EA) category. GA is commonly used for high-quality solutions and problem optimization (Albadr et al., 2020). Evolutionary algorithms are commonly used to solve problems that do not have a well-defined or efficient solution. This approach can solve optima, near optima, shortest path, or schedule. It solves the problems by generating random sets of solutions, also known as candidates. These represent possible solutions to the problem that are described by a list of characteristics or genes that will be changing or evolving with each iteration (Alam et al., 2020).

Once the population has been initialized, the GA-based optimization process typically follows a four-step procedure:

1. Calculate or evaluate fitness, here it calculates how likely this solution is to be chosen, or how well this candidate solves the problem.
2. Selection process, where solutions are compared to one another to draw a set that will function as the parent for the next generation.
3. Reproduction is the step in which previously chosen solutions are paired to form offspring. This is achieved using crossover, where genes are combined, and mutation occurs to provide diversity and explore new solutions.

4. Lastly, termination, where the algorithm checks a certain criterion to decide whether to continue with the next generation or whether the process can end. This can be achieved by reaching a certain score when evaluating or after some set number of generations.

The proposed model (figure 2) is based on the previously explained process of genetic algorithms. It is mainly defined by 6 major components: data cleaning and preprocessing, genetic algorithm layout, random solutions, evaluation, selection, and reproduction. It is important to mention that some errors were observed within some spatial analysis tools when ArcPy (Python site package) is used outside of ArcGIS Pro. The model was built within the ArcGIS Pro notebook.

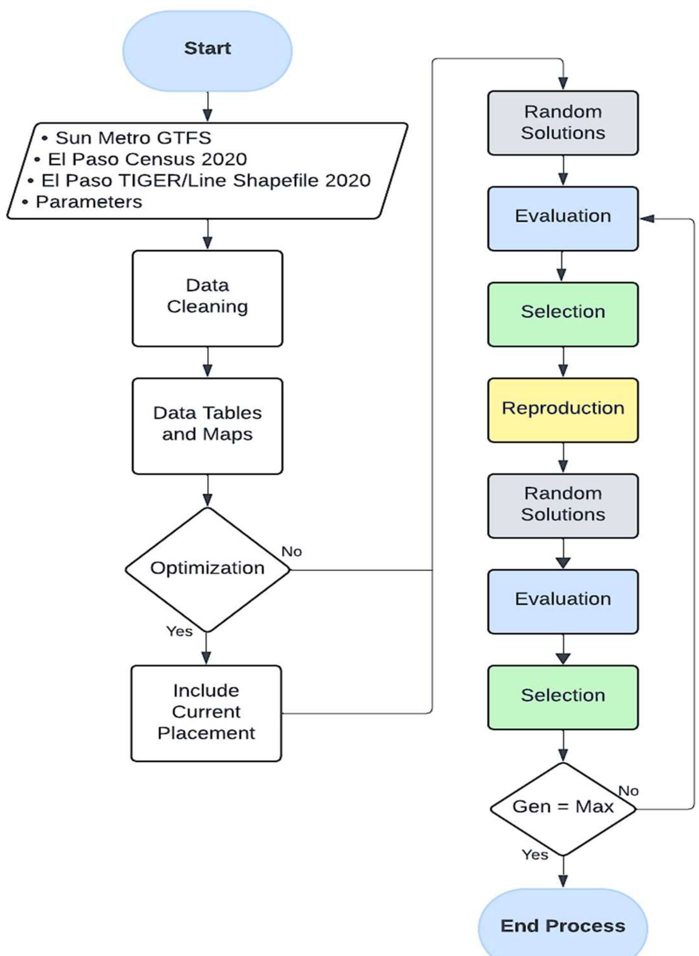

Figure 2. Methodology

DATA CLEANING AND PREPROCESSING

In any research, the quality of the analysis is entirely dependent on the quality of the data used. Therefore, meticulous attention to data cleaning and preprocessing is key to ensuring the accuracy and reliability of the results. The following section explains the steps taken to prepare the raw data. Starting with data cleaning, addressing common challenges of working with TIGER files and census data, and emphasizing correct data selection to maintain some level of data integrity. Followed by the preprocessing steps, which include creating tables and working with GTFS data. This step ensures the data is in the appropriate format and consistent, setting a strong foundation for the analysis.

Data Cleaning

When working with TIGER files and census data, it is important to download data from the same year or close to the same year since census divisions can change from one year to another. This difference in census divisions can cause problems such as zones sharing the same GEOID and indexes without a match (Table-Map). The best approach is to base your data selection on the closest decennial survey or American community survey.

It is common to see census blocks sharing the same GEOID and the same information. When calculating indicators or demographics, this can cause the numbers to be highly inaccurate since the model will add the same population more than once for multiple zones. A recommendation that will assist in identifying and solving the problem is to summarize the tables when exporting the census data to the workspace in ArcGIS PRO. To set up this tool, select the GEOID (CB) as the case field, and then, for the statistics field, select the data with the option of First. The resulting tables will have the GEOID, the frequency of the GEOID, which is just the number of appearances of that GEOID in the table, and the first value in the data for that GEOID. This will prevent repetition when using spatial analysis tools.

Repetitive GEOIDs in the shapefiles are not always an issue since this model only uses the maps for area calculation, spatial analysis, and visualization while running the evaluations in the tables. In the case of needing to work with the maps and data simultaneously, dissolving the map by GEOID will combine all the features or zones that share the same GEOID, this will change the appearance of the map but will prevent duplicating information.

Data Preprocessing

Data preprocessing is a crucial step, as it involves preparing and transforming raw data into readable formats for the model to use. The following sections detail the modifications and processes that tables, maps, and features undergo:

Data Calculation Table

Most non-spatial calculations will use this table. The information that will be required in this table is the demographics of the zone or areas to study at the census block level. After cleaning the data, the only requirement for information already available at the CB level is to use the join field function to create a single table containing the demographics. For data not available at the census block level, a transformation is required. First, it is necessary to create a Census Block Group; this table will serve as the calculation table to transform the data. This table will contain the total population by GEOID (CBG) and the demographics that are not available at the CB level. From this table, a demographic-to-population ratio can be obtained; the ratio shows the likelihood of a member of a certain demographic being at a CB based on the total population. This is possible by assuming that all CBs within the same CBG share the same ratio. With this information, the data calculation table can be completed just by multiplying the total population of each CB by its

CBG demographic ratio. One of the limitations of this approach is that it only works when transforming data from CBG to CB information from a higher level, such as census tracts or zip codes, which requires more assumptions, reducing the accuracy of the analysis.

GTFS

The transit system information is available within the GTFS, as it contains details in the form of text files, shapes (bus routes), stops, schedules, or trips. From this dataset, the required files are shapes and bus stops. With ArcGIS Pro converting these text files, there are two features: a set of line features that describe the route for each bus route, colored and named differently to be easily identified. The second feature contains sets of X and Y points that locate each available bus stop served by the current transit system.

To continue with the route selection, it utilizes the attribute table of the routes, which contains a field with the route number, to select the correct route. As the bus stops do not have a direct correlation to the routes, the option is to utilize selection by location on ArcGIS Pro to select what could be the correct set of stops that are on top of the route or in proximity to it.

Shapefiles

The selection process of the maps (census shapefiles) is also based on the selected route, just like the selection of the bus stops. It is important to know the effective range of accessibility, in other words, how it will measure accessibility. In the proposed model, the recommended walking distance to a bus stop is set at one-quarter of a mile or four hundred meters. With this value, the CBs that are within that distance were selected. This selection is important to gain the correct insights into how the changes in bus stop locations affect the accessibility of the area surrounding area.

Genetic Algorithm Structure

The GA employed in this study was designed to optimize the spatial distribution of bus stops along the given route. Using an iterative process, it refines the pool of potential solutions, enhancing fitness with each generation. This decision-making process hinges on the accessibility of designated demographic groups to the transportation system. Here is an overview of the key steps involved:

Steps 1 & 5: Initialization – Random Solutions

The model generates the initial set of solutions to evaluate the first generation of candidates during the initialization step (Figure 3). It is based on a list of parameters that will guide it in generating solutions, including population size, mutation rate, route location, and proximity constraint. This initial or first generation will either be entirely random solutions or will contain the original placement if working with problem optimization.

The number of solutions is required because it indicates the size of each generation. This number will be used to determine when to finalize the random solution generation process. The number of solutions can vary depending on the objective, evaluation function, computational capacity, other constraints, or study requirements. Another parameter required is the desired number of bus stops. It can be the same as the original solution placement or a different number if required.

Since the function requires coordinates or any other type of geospatial reference to locate the generation of random solutions, it uses the line feature (route) provided to get a location on which these random solutions will be based. Now, it is important to note that bus stops will not be exactly on top of the route; providing a maximum distance to the route gives a more realistic approach to the placement of bus stops. Due to computational constraints, analyzing a vast number of solutions per generation was not possible. A small population size can make the model fall into a local minimum due to the limited options or diversity. To prevent this, a new set of solutions is introduced to provide fresh options for each generation. This function will be called again right after the reproduction process. The difference this time is that now it will not be called to generate an initial population but new sets. The new random solutions will have the same parameters as the initial ones but will have a smaller population size.

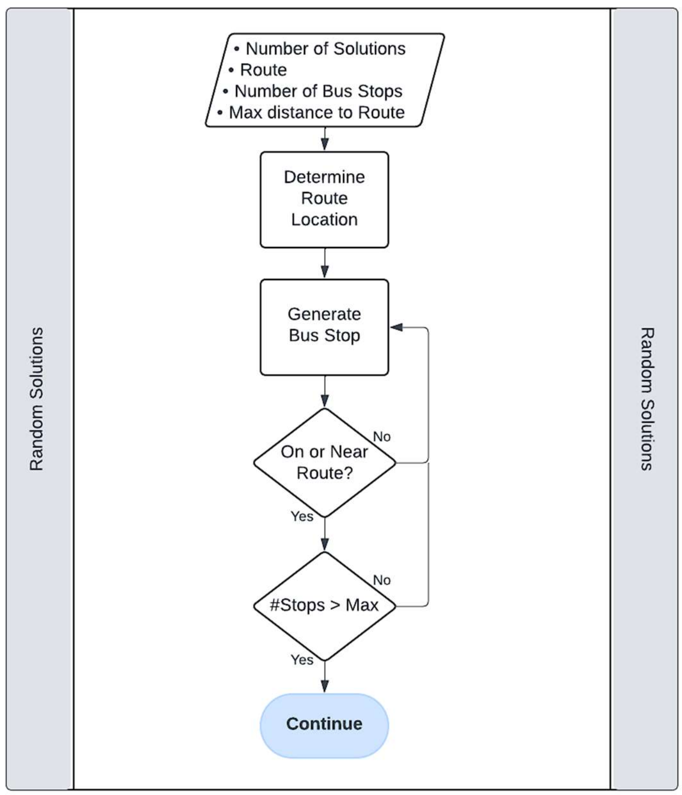

Figure 3. GA Random Solutions

Steps 2 & 6: Fitness Evaluation

Genetic algorithms base the selection process on the fitness score of the solutions, which weighs the capacity of that candidate to satisfy the objectives or meet the constraints of the given problem. This value will later be used during the selection process for parenting/reproduction and to determine which solutions are suited to survive and contribute to the next generation. It is important to note that this function only evaluates one solution at a time.

Figure 4 shows the structure used to evaluate these solutions, which is a 5-step process. The inputs to run this analysis are the dictionary containing the solutions with its own set of coordinates, data maps, which are the maps created during the data preprocessing, the data calculation table, and the buffer distance.

Since the solutions are stored as a set of X and Y coordinates, the function will transform them into a point feature to be later used. Based on this newly created point feature, it will create buffer zones based on the specified distance (walking distance). When setting the use of a buffer in ArcPy, it is crucial to select dissolve “ALL” to create a single feature instead of multiple buffers. Using the buffer zones and the data map intersecting these two features will allow the model to calculate the coverage area of each census block. The percentage of accessibility for each demography can be calculated based on the coverage information. These values will be used as the fitness score during the selection process. Once completed, the process will continue until all solutions have been evaluated.

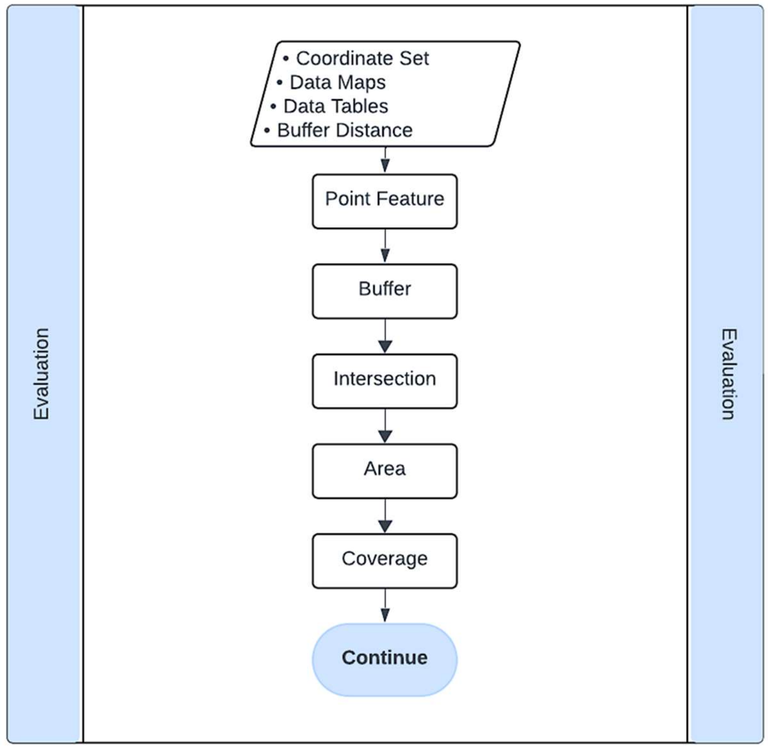

Figure 4. Fitness Evaluation

Steps 3 & 7: Selection – Roulette Wheel Selection

The selection process is one of the most important steps towards funding the best solutions using a GA. There is more than one method that can be used to serve this function, but for this model, the chosen selection technique is Roulette Wheel Selection (RWS). Utilizing Roulette Wheel Selection (RWS) offers several advantages. Firstly, it ensures unbiased selection as all solutions are given an opportunity, thereby maintaining diversity within the population. Additionally, it helps prevent any single candidate from gaining undue influence or control.

Two disadvantages that can be a concern are that it can be a time-consuming process due to the ranking nature of this technique and that if solutions converge earlier than anticipated, the method may lose its efficiency.

There are two main inputs to initiate the selection. The maximum number of solutions provides the number of candidates that will either be selected for parenting and reproduction or the candidates that will survive for the next generation. The second input is the scores for each solution. The first step in the structure of the selection process and RWS (Figure 5) consists of the summation of all scores to calculate total fitness. This value is then used to normalize scores, which is achieved by dividing the scores by the total fitness. A roulette wheel is then designed, in which everyone is given a portion based on their normalized fitness score. The final selection is done by generating a random number between 0 and 1, and the candidates that meet that probability will then be chosen.

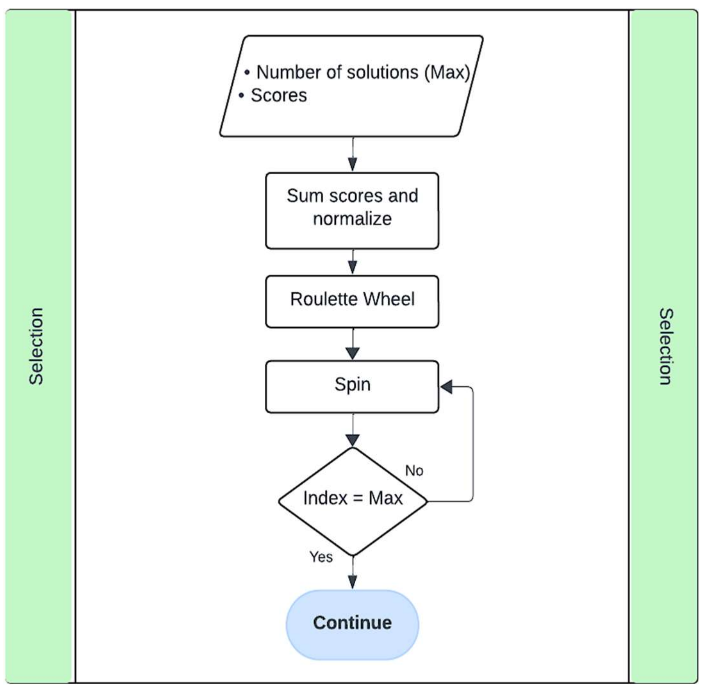

Figure 5. Roulette Wheel Selection (RWS)

Step 4: Reproduction

The reproduction function is based on genetic operations, like crossover and mutation with the selected candidates to form the offspring (figure 6). It starts by pairing the selected solutions;

a recommendation is to select an even number of candidates for parenting to prevent errors. Then, following the pairing of the selected candidates, the crossover function or recombination is performed. The idea behind it is to exchange genetic material between the chromosomes of selected parents. In this case, each gene represents one of the bus stops of each solution. This will form offspring that will be based on the genes (bus stops) of the parents.

Now, the mutation process is used to provide diversity and modify the genes of each solution. The function assigns a number between 0 and 1 to each gene, and if that number is lower than the mutation rate, slight modifications or mutations will be introduced to the gene.

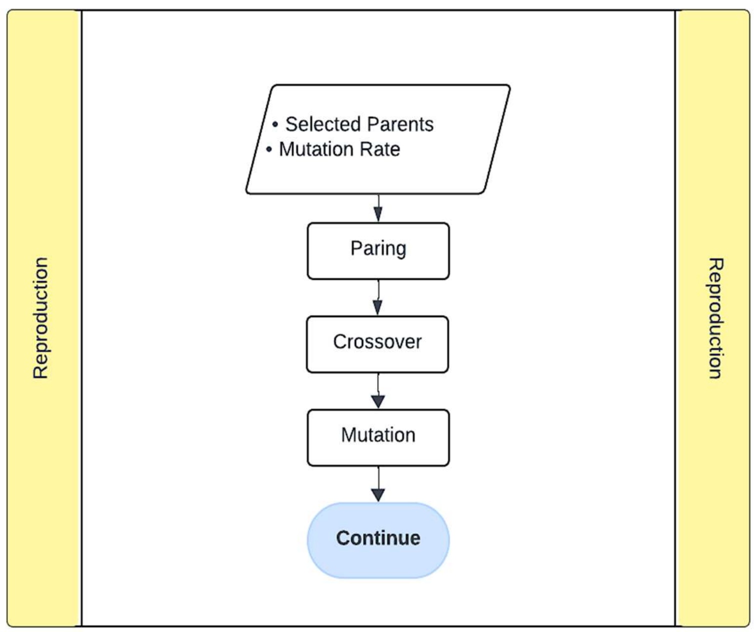

Figure 6. Reproduction

DATA LIMITATIONS AND PERFORMANCE

Gathering data poses a common challenge in transportation planning. In this model, for instance, publicly accessible data sourced from the Census Bureau and transit agencies is utilized. While this data is highly accurate, its scope is often constrained by the specific needs or available resources of these entities. Conversely, private companies specializing in data collection maintain extensive repositories online but behind a paywall, which restricts accessibility to a wider audience.

Cultural, economic, and geographic factors pose challenges when gathering or updating accurate information. El Paso encounters several of these challenges, which can limit the precision and scope of this study. Firstly, the datasets utilized are from the year 2020, coinciding with the onset of the Covid-19 pandemic. During this unprecedented period and in the subsequent years, El Paso, like many other U.S. cities, faced various challenges. These included sudden spikes in unemployment and a transient population influx as individuals from other cities sought the city's lower cost of living and some locals relocated to Ciudad Juarez to be closer to family and mitigate expenses.

In recent years, El Paso has been facing an immigration crisis, as many people from Central and South America are trying to cross the border into the US to increase living conditions and have better opportunities. As this situation is moving from an extraordinary condition to a normal characteristic, transit planning needs to take this information into account. Most of the time, these communities would fall under disadvantaged groups, which should be considered in this study. However, the lack of information, as they are a transit population, makes it hard to account for them, reducing the overall accuracy and effectiveness.

As mentioned previously, this study utilizes information from census blocks or census block groups. This drastically restricts the extent of this study as it is limited to the available data since using data at other geographic levels, such as Census tracts, Zip codes, or TAZs, limits the accuracy of the model.

Something that needs to be considered when using genetic algorithms and GIS-based solutions is performance, such as the accuracy of results, cycle duration, computational demand,

and error handling. The model constantly provided solutions or candidates within the specified spatial parameters, such as proximity to the route and area selection based on the walking distance. Calculations of areas and percentages are consistent through the generations, giving few to no issues.

Managing cycle duration presented a challenge due to the substantial computational requirements associated with working in ArcGIS Pro and employing Genetic Algorithms (GA). This study utilized two machines (Table 4). The primary hurdle in optimizing cycle time is in managing RAM saturation, particularly with the machine equipped with an AMD GPU. This was exacerbated by the fact that ArcGIS Pro exclusively supports Nvidia GPUs, necessitating reliance on regular RAM to execute all processes.

Table 4. Machine Specifications

	ROG Zephyrus G14	Alienware
CPU	AMD Ryzen 9 6900HS	Intel i7
GPU	AMD Radeon RX 6700S	Nvidia GTX 1080
RAM	16 GB	32 GB
VRAM	16 GB	24 GB
Cycle Time	25 seconds	55 seconds

Chapter 4: Data Analysis

TRANSPORTATION ACCESSIBILITY IN THE COUNTY OF EL PASO, TX

The County of El Paso in Texas has been selected to run a case study for this research. To gain a deeper understanding of the analysis within the selected area, it's crucial to understand the current transportation accessibility situation at the City/County level. Although this study focuses on micro-level analysis (one bus route), in the actual development process of a public transportation system, understanding the overall accessibility is crucial. This serves as the initial step in assessing more specific zones and understanding the "general" accessibility, which aids in identifying existing equity and accessibility issues such as transit deserts, inefficient routes, or underserved demographics. These considerations should always be prioritized when planning updates to the transit system or selecting new developments. Moreover, this preliminary analysis can establish a benchmark—a reference point for future comparisons. For instance, objectives can be set to address specific Census Block Groups (CBGs) to align with the current accessibility at the city level.

The proposed model, or a segment thereof, can be adapted to analyze the existing conditions at the city level. Specifically, the fitness evaluation function of the model can be repurposed for this task. Leveraging this particular section is advantageous because it ensures that comparisons to the actual solution or metrics are grounded in the same scoring system and parameters. To achieve this, some modifications should be made first. Creating a functional model that only evaluates accessibility based on an equitable criterion requires copying the previously mentioned function with all the parameters, variables, and subfunctions.

According to the decennial census of 2020, El Paso County has a total population of approximately 865,000. Its population is characterized by an overwhelming 82% of Hispanics (Table 5), which contrasts with many cities in the United States. However, this proportion aligns closely with other cities in Texas, Florida, and California—states with similar cultural heritage and demographic conditions characterized by significant Hispanic influence, primarily from Mexico and Cuba. White population represents 11.2% of the total population, a high percentage compared to African Americans and Asians, with 3.36% and 1.39% respectively (U.S. Census Bureau, 2023). The preliminary accessibility analysis (quarter-mile buffer around bus stops) for various populations is shown in Table 5.

Table 5. Demographics and Accessibility (%)

	Total	Percentage	Population Served	Accessibility (%)
General Population	865657	-	430718	49.76
White	96953	11.2	48059.89	49.57
Hispanic	715351	82.64	361535	50.54
Black/African American	29054	3.36	14135	48.65
Asian	12073	1.39	5284	43.77
Poverty	160122	18.50	91828	57.35
Total Labor Force	378287	43.70	198243	52.41
Labor Force Unemployed	23499	2.71	13006	55.35

Analysis has shown that around 50% of the general population enjoys accessibility to the transit system, leaving the remaining half without access, as illustrated in Figure 7. This figure shows regions like Lower Valley, Far East, Canutillo, and Upper Valley, where high-density areas coincide with low accessibility to transit services.

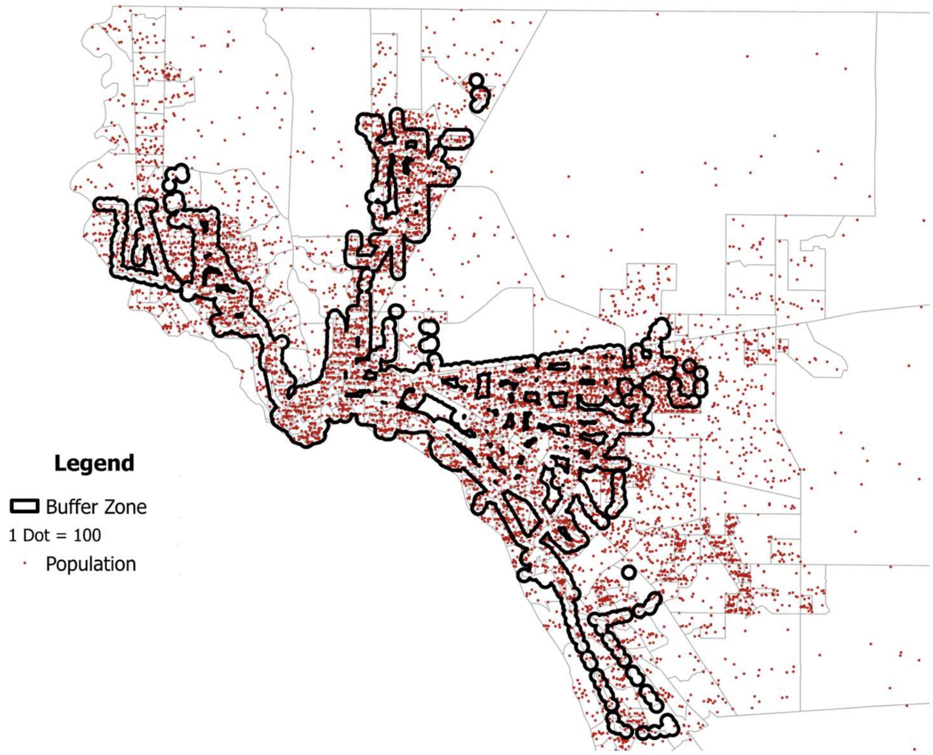

Figure 7. Dot Density Map with a Quarter Mile Buffer

The overall percentage accessibility is close to 50% for all the demographics that are taken into consideration for this analysis. The areas with an Asian population in El Paso present an interesting pattern; despite being the smallest demographic percentage-wise, they also have the lowest accessibility percentage. In the dot density map shown in Figure 8, it becomes apparent that most of the Asian population lacking coverage resides predominantly in Fort Bliss or West Side El Paso. Notably, Upper Mesa Hills and Northwest El Paso, segments of the West Side, are areas populated by medium- to high-income households, which could account for their notably limited accessibility.

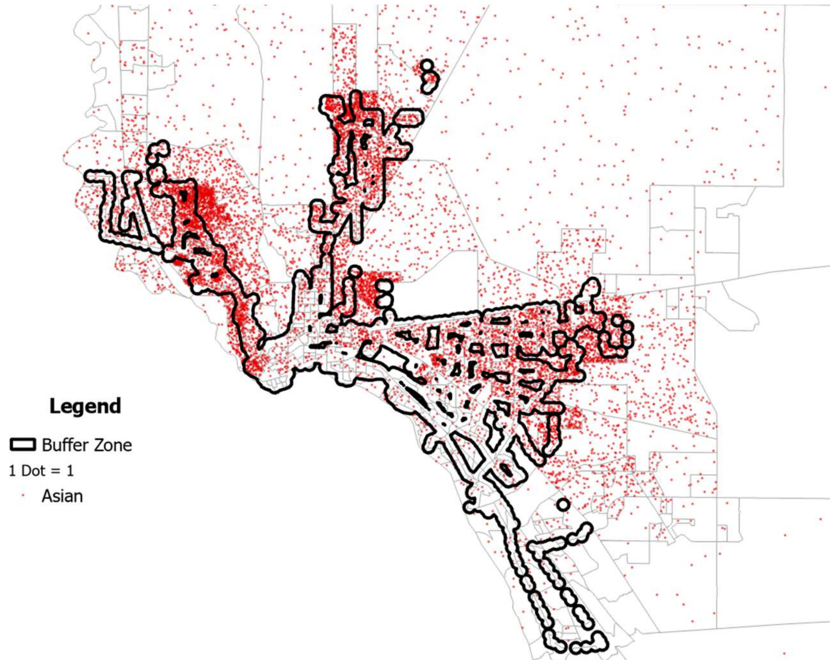

Figure 8. Dot Density (Asian) with Buffer Zone

Figure 9 shows maps illustrating the percentage of accessibility across various demographics. Upon comparing the tabulated results to the mapped representations, a striking similarity emerges: both exhibit analogous percentages of accessibility, resulting in visually comparable maps.

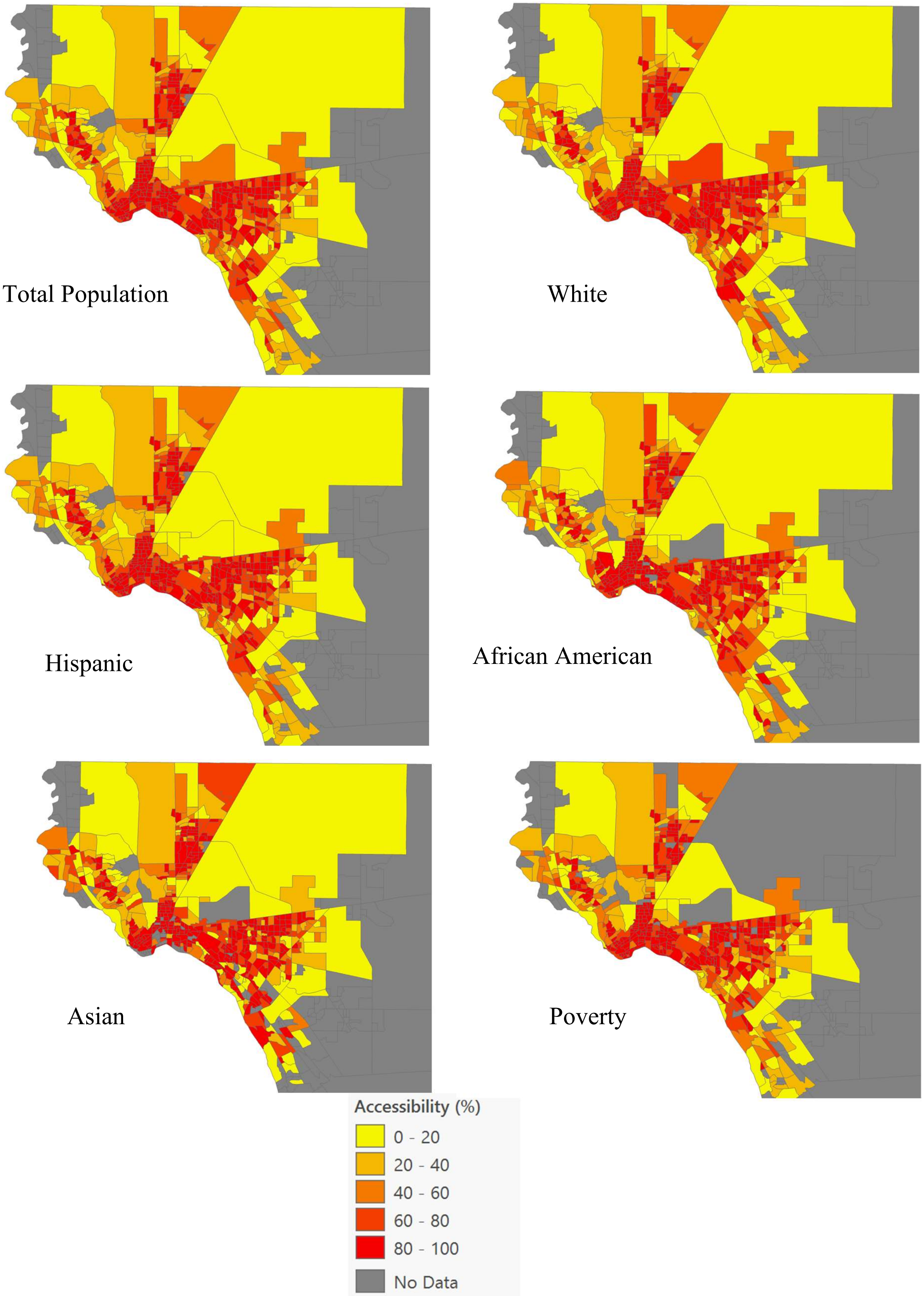

Figure 9: Transportation Accessibility by Percentage

ROUTE 16 & SURROUNDING AREAS

The selected route for this study is Route 16, also known as Upper Valley, a route classified by Sun Metro as a feeder connecting the Westside Transit Center to the neighborhoods located north of Country Club Rd and west of Doniphan Dr (Figure 11). With a total of thirty-three bus stops, the main arteries that it runs through are Country Club Rd, Upper Valley Rd, Artcraft Rd, Doniphan Dr, and Montoya Dr (Figure 10). According to Sun Metros State of the System Report (2022), it only operates on weekdays, with just one vehicle assigned to the route; it has the lowest ridership of any route, mainly due to its very limited schedule.

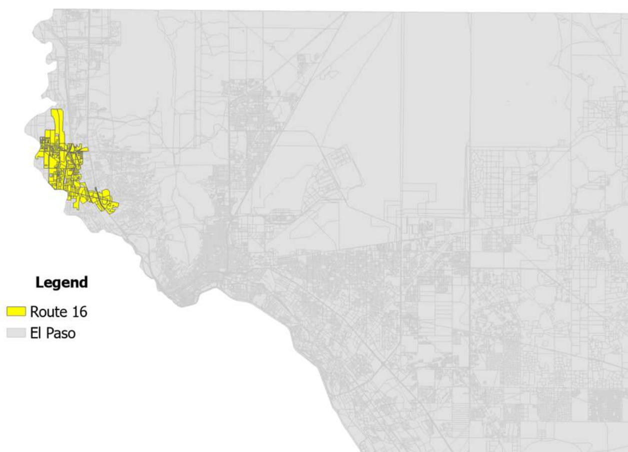

Figure 10. Route 16 Spatial Reference

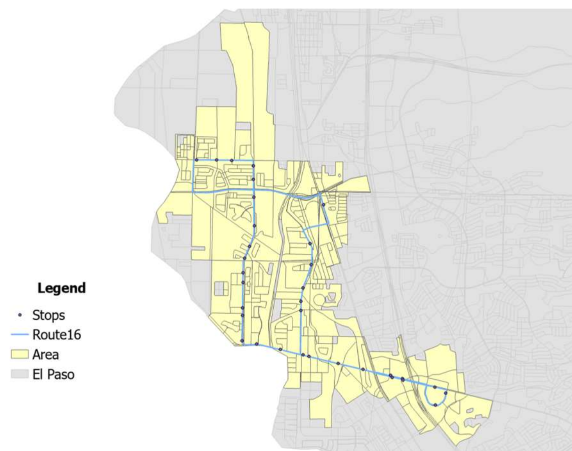

Figure 11. Route 16 and Stops

The surrounding area comprises all census blocks with access to the route within a 400-meter walking distance radius. This area, which will be referred to as "Upper Valley," has some differences in its demographic composition compared to El Paso County (Table 6). For instance, the white population has 5% more presence in this region; another notable difference is the presence of Hispanics, who have 7% less presence than the County. Both poverty and unemployment levels are lower, indicating that the community has better economic conditions when compared to the County.

Upon examination of accessibility levels, it's evident that there are significantly lower rates in this area. For example, the average accessibility dropped from 51% to 38%. Reductions span from 4% for the White population to as low as 28% for individuals in poverty. It is a region with much higher economic conditions, but that should not justify the poor transit conditions and marginalization of local disadvantaged communities. Unfortunately, this does not provide the whole picture of "Upper Valley" since it only considers the areas with potential access to the current transit system and omits other accessibility factors like schedule or connectivity. There are more CBs within "Upper Valley" that have 0% accessibility since they do not have a route close enough to them.

This analysis shows the need for updates and improvements in the current transit system within this region. To address the need, this study proposes a hybrid model by utilizing Genetic Algorithms (GA) in conjunction with Geographic Information Systems (GIS) to either optimize the current layout of bus stops or suggest a new solution or arrangement based on the equity factors previously outlined.

Table 6. Accessibility (%) for Route 16

	Total	Percentage	Population Served	Accessibility (%)
General Population	43652	-	19544	44.77
White	17895	40.99	8121	45.38
Hispanic	32547	74.56	14235	43.74
Black/African American	876	2.01	331	37.84
Asian	646	1.48	230	35.53
Poverty	5862	13.43	1735	29.60
Total Labor Force	20564	47.11	8720	42.41
Labor Force Unemployed	1072	2.45	331	30.87

OPTIMIZATION PROBLEM

For optimizing the current placement of bus stops, the final solutions should be rooted in the original set of coordinates. This approach anticipates observing the genetic traits of the original set passed down to subsequent generations under the assumption that the current placement represents a sound or nearly optimal solution based on the established fitness evaluation parameters. However, one of the challenges identified during data preparation was the initially low accessibility inherent in the current layout. This is concerning, as the model might inadvertently discard it in the initial iterations of the process, rendering the model unsuitable for this dataset.

Table 7 provides a summary of the percentage accessibility and the total fitness score. Here, we have information from the original set (0), 10 random generations (1-10), and the best solutions after the last generation (11-13). It is important to remember that this study prioritizes selecting optimal solutions with equity considerations at the forefront. The score used in the selection process was determined based on the accessibility of Black, Asian, and economically disadvantaged individuals.

The original set got a score of 102.97, with African Americans accounting for 38%, Asians for 35%, and the economically disadvantaged population for 29.6%. These are low percentages compared to the average in El Paso. The first generation, instead of improving accessibility, dropped it drastically, reducing scores to 75.

This behavior is typical of Genetic Algorithms, wherein the initial generations explore various solutions, often encountering less favorable solutions along the way. However, as generations progress, accessibility and the overall score gradually improve (Figure 12). In this approach, African Americans and Asians experience the most significant increase in accessibility, with percentages rising by approximately 2% and 5%, respectively, while the poverty-stricken population sees no change. This outcome arises from assigning equal weight to all demographics, implying that each holds the same value and contributes equally to the selection process.

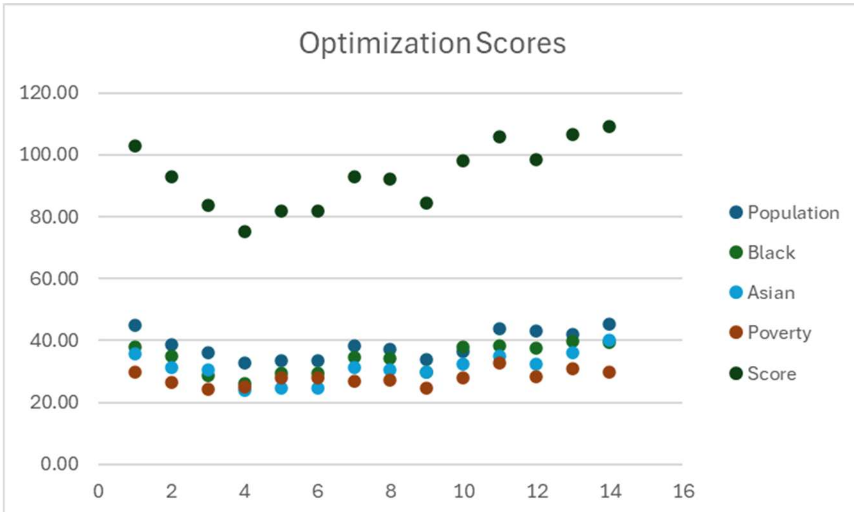

Figure 12. Optimization Scores

Figure 13 presents 4 maps. The original solution, with a score of 102.97, was the main distribution of its bus stops along Country Club Rd and Upper Valley Rd, with no stops on Artcraft. Solution 3, with a score of 75.07, is a good example of an error that can happen; on this occasion, two bus stops are even outside the selected area. Solution 9, with a score of 98.14, presents a higher stop density along Country Club, which is understandable since it is a major road surrounded by neighborhoods, but it just has too many for that single road. Solution 13, with a score of 109.28, presents some of the genes left by the original solution, the difference being that this solution distributes its bus stops more evenly around all roads, including Artcraft.

Table 4. Accessibility (%) and Scores

	Population	White	Black	Asian	Hispanic	Total Labor Force	Labor Force Unemployed	Poverty	Score
0	44.77	45.38	37.84	35.53	43.74	42.41	30.87	29.60	102.97
1	38.79	38.89	34.96	31.37	38.93	36.42	26.84	26.60	92.94
2	35.98	34.83	28.52	30.63	36.81	33.74	23.94	24.39	83.54
3	32.75	32.21	26.21	23.84	33.10	30.57	22.60	25.01	75.07
4	33.47	35.43	29.36	24.51	32.04	32.91	24.04	27.99	81.86
5	33.47	35.43	29.36	24.51	32.04	32.91	24.04	27.99	81.86
6	38.11	38.05	34.70	31.32	38.04	35.94	24.76	26.80	92.81
7	37.03	38.95	34.28	30.68	35.73	36.12	20.42	27.13	92.09
8	33.77	36.05	29.88	29.73	32.31	32.96	16.26	24.69	84.30
9	36.56	37.37	37.96	32.19	35.54	35.73	26.29	27.99	98.14
10	43.77	43.42	38.18	35.06	43.45	41.67	33.85	32.75	105.99
11	43.14	42.45	37.36	32.55	43.25	40.65	27.64	28.37	98.29
12	41.97	41.44	39.62	36.24	42.30	40.12	25.73	30.82	106.68
13	45.22	44.69	39.42	40.22	45.45	42.56	33.94	29.64	109.2

Figure 13. Optimization Solutions

NEW SOLUTION PROBLEM

The new solution problem refers to those cases in which the transit planner knows the possible trajectory or path that a proposed route will take and needs to design for its bus stops. When designing new bus routes, it is important to take into consideration the needs of the residents. This model is designed to recommend potential bus stop locations, with a focus on maximizing equity factors (demographic indicators) as the primary variable. To evaluate the model's capability in suggesting bus stop locations, route 16 is utilized, with existing bus stop locations serving as a benchmark for result comparison. Table 8 displays the accessibility percentages for each demographic within specific solutions generated across random generations. Table 8 includes the original set (0), which serves as a reference without impacting the solutions, the top solutions from random generations (1-10), and the two best final solutions (11-12). The numbering on the left of the table and along the x-axis of the graphs doesn't correspond to actual generations; instead, it provides chronological reference points.

Table 8. New Solution Results

	Population	White	Black	Asian	Hispanic	Total Labor Force	Labor Force Unemployed	Poverty	Score
0	44.77	45.38	37.84	35.53	43.74	42.41	30.87	29.60	102.97
1	23.32	25.12	22.87	20.99	21.82	22.95	8.37	15.74	59.61
2	28.12	30.75	26.22	23.25	26.37	27.77	11.01	17.72	67.19
3	25.08	27.42	25.52	21.49	23.26	24.49	10.92	16.67	63.69
4	27.92	29.81	29.27	26.38	26.49	27.70	14.52	21.19	76.84
5	39.97	41.40	32.94	32.44	38.73	37.67	23.27	24.16	89.54
6	42.52	43.59	34.75	33.76	41.67	40.78	24.01	28.12	96.63
7	36.50	38.84	33.58	32.78	35.08	35.78	22.81	25.86	92.22
8	37.53	37.65	35.15	32.12	37.86	36.11	24.17	25.97	93.24
9	39.54	39.02	37.11	33.33	40.14	38.43	24.47	28.96	99.40
10	40.91	39.99	33.98	33.47	41.54	38.61	26.25	30.69	98.13
11	43.38	42.57	38.64	33.69	43.68	41.05	27.28	29.02	101.35
12	41.82	41.58	39.64	36.13	42.00	39.76	26.04	31.04	106.82

In this example, we observe marginal increases in accessibility, ranging from approximately 1 to 2 percent per demographic. While these improvements may seem modest at the micro level, when integrated into the entire system, they could potentially enhance accessibility for thousands of individuals. Additionally, the analysis indicates that even upon completion of the run, the candidates had yet to exhibit signs of convergence, suggesting that the model may need either additional generations or a larger pool of potential solutions. The solution with the highest index resembles the original placement, suggesting it's heading in the right direction. However, notable disparities exist between the original set and the top two solutions, particularly evident at two crucial points: West Side Dr and Artcraft, and near Doniphan Dr.

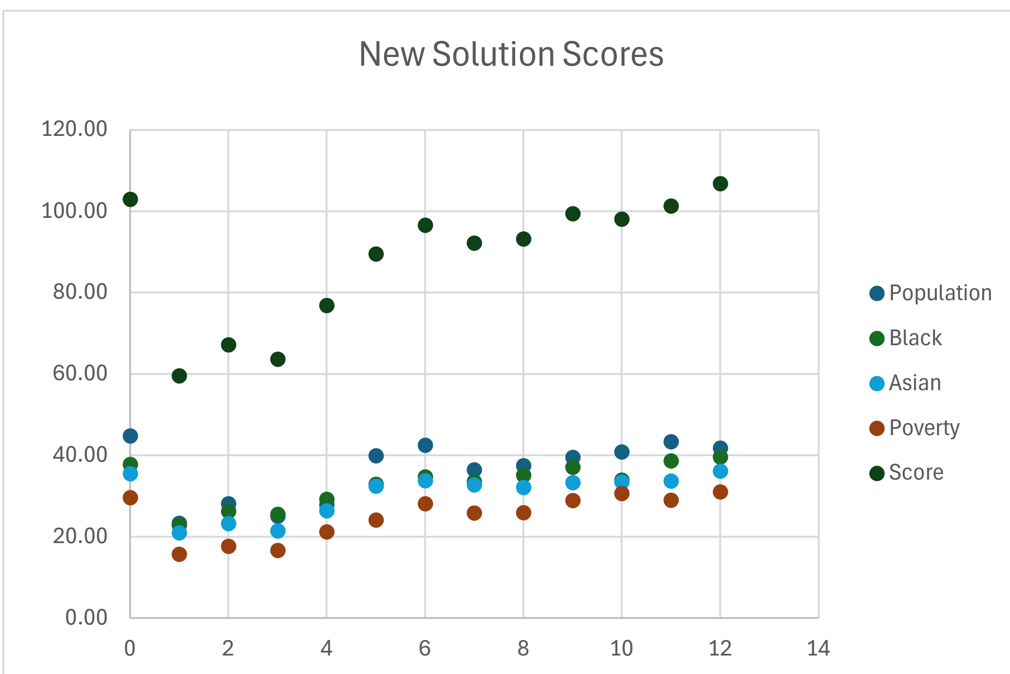

Figure 14. Accessibility (%) and Scores

Figure 15. Optimal Solutions

POTENTIAL IMPACT ON RIDERSHIP

The potential impact of optimization in terms of users can be determined by calculating the accessibility percentages for each demographic group and the general population. This figure is crucial for understanding the effects of these adjustments on both the transit system and the population. Table 9 illustrates the best-case scenario for the increase in potential users or the enhancement in accessibility for each demographic group. For instance, assuming an average increase of 1.8% in ridership, considering that minorities and individuals in poverty constitute the primary users of the public transit system, the ridership data is extracted from Sun Metro's State of System Report.

Table 5. Potential Users and Ridership (Upper Valley)

Demographic	Users	Percentage Increase	Potential Increase In users
Black/African American	876	1.8%	16
Asian	646	4.69%	30
Poverty	5862	1.44%	84
Total Population	43625	0.45%	196
Ridership Annually	9360	1.8%	168

If now the same increase of accessibility is assumed for the whole transportation system in El Paso, the ridership numbers would be as follows.

Table 6. Potential Users and Ridership (El Paso)

Demographic	Users	Percentage Increase	Potential Increase
Black/African American	29054	1.8%	523
Asian	12073	4.69%	566
Poverty	160122	1.44%	2305
Total Population	865657	0.45%	3895
Ridership Annually	5855200	1.8%	105393

Chapter 5: Conclusion & Future Work

Social justice has long been a concern, but it wasn't until recent years that society truly grasped the extent of the problem. Over the years, incomplete and outdated transportation planning methodologies have marginalized certain communities, often without recognition. Progress and urban development should never compromise the quality of life in communities. Yet, year after year, disadvantaged communities bore the brunt of transportation projects that didn't account for their needs. Despite advancements with initiatives like the Justice40 Initiative, which mandate greater resource allocation to underserved communities or projects that benefit them, there remains a lack of clear metrics for measuring equity. Standardizing parameters and offering frameworks could be the key for agencies and urban planners to enhance the quality of life in these communities.

The integration of Genetic Algorithms (GA) with Geographic Information Systems (GIS) presents a promising approach to address equity concerns while increasing accessibility in future transportation projects. By leveraging the capabilities of these technologies, transportation planners can analyze complex spatial datasets, providing the tools to optimize transit systems while ensuring equitable access to surrounding communities. Through the iterative process of GA, planners can identify the optimal or near-optimal locations for bus stops that will prioritize equitable access to the transit system.

Data availability and quality are major barriers to the successful implementation of transportation planning strategies that prioritize equitable access. Although the used datasets are reliable sources of information, the precision of the results can be affected by the level of detail and accuracy. This study employed GA and GIS tools to optimize the placement of bus stops with a focus on ensuring equitable access to transportation systems. The analysis was carried out using the City of El Paso as a case study, with Sun Metro’s Route 16 chosen for microanalysis. The findings revealed disparities in transit accessibility, particularly in the Upper Valley area along the route. Despite relatively favorable economic conditions, residents in this area face significant gaps in the existing transit system. Disadvantaged communities are especially affected, highlighting the need for immediate action from local authorities and government agencies.

The deployment of the suggested hybrid model for optimizing existing bus stop locations demonstrated its ability to enhance current placements while addressing equity issues. Although the model's best solution only results in a 2 percent increase in accessibility for disadvantaged communities, the upward trend of the scores in the graph (Figure 12) indicates that it has not yet converged to a solution. This suggests that further improvements could be achieved by either increasing the number of generations or the number of solutions, contingent upon enhanced computational capabilities.

When the model is configured to discover new solutions for novel bus routes, the initial solutions exhibit notably poor performance. It's not until later generations that accessibility begins to improve. Like the optimization challenge, in the final generation, the percentage increase in accessibility compared to the current placement is in single-digit figures. While this may appear modest, it's crucial to acknowledge that the existing placement likely prioritized service provision to the majority rather than equity considerations. The results show again a positive slope in the last generations, indicating that with more time or options (generations or number of solutions), it could have provided even better solutions.

Based on the limitations and results obtained from this study, there are a couple of directions for improvement in the continuation of this project. Here's a brief list of potential modifications, improvements, or ideas that could be explored in the future:

1. Data has been one limitation. Collecting more data on demographics and other indicators like travel time, travel trends, and travel purposes could increase the accuracy of the proposed model.
2. Study the possibility of integrating time, cost, or connectivity. This will require a more comprehensive fitness evaluation function, but the results could potentially be more realistic.
3. Optimizing the algorithm to reduce the cycle time is crucial. A solution could be transferring the workload from the CPU to the GPU, as done in other studies with good results (Aqib et al., 2019).
4. Lastly, community engagement is crucial for gathering insights into the local population's needs and preferences to create the best model possible. It is impossible to solve problems that you do not understand.

In conclusion, proactive transportation planning strategies are essential for tackling society's challenges and ensuring equitable access to services for all, not just a select few. Leveraging the power and capabilities of new technologies such as GA and GIS holds promise for addressing these challenges and equity issues in transportation. We have progressed as a society to have the necessary means and technologies to solve this problem; it's our responsibility to utilize them effectively and make a meaningful difference, particularly for historically underserved communities.

References

Abbasi, M., Rafiee, M., Khosravi, M. R., Jolfaei, A., Menon, V. G., & Koushyar, J. M. (2020). An efficient parallel genetic algorithm solution for vehicle routing problem in cloud implementation of the intelligent transportation systems. *Journal of Cloud Computing*, *9*.

Abu-Lebdeh, G., Chen, H., & Ghanim, M. (2016). Improving Performance of Genetic Algorithms for Transportation Systems: Case of Parallel Genetic Algorithms. *Journal of Infrastructure Systems*, *22*(4). https://doi.org/10.1061/(ASCE)IS.1943-555X.0000206

Adli, S. N., & Chowdhury, S. (2021). A Critical Review of Social Justice Theories in Public Transit Planning. *Sustainability*, *13*(8), 4289. https://doi.org/10.3390/su13084289

Alam, T., Qamar, S., Dixit, A., & Benaida, M. (2020). Genetic Algorithm: Reviews, Implementations, and Applications. *International Journal of Engineering Pedagogy*.

Albadr, M. A., Tiun, S., Ayob, M., & AL-Dhief, F. (2020). Genetic Algorithm Based on Natural Selection Theory for Optimization Problems. *Symmetry*, *12*(11), 1758. https://doi.org/10.3390/sym12111758

Antipova, A., Sultana, S., Hu, Y., & Rhudy, J. P. (2020). Accessibility and transportation equity. In *Sustainability (Switzerland)* (Vol. 12, Issue 9). MDPI. https://doi.org/10.3390/SU12093611

Aqib, M., Mehmood, R., Alzahrani, A., Katib, I., Albeshri, A., & Altowaijri, S. M. (2019). Rapid Transit Systems: Smarter Urban Planning Using Big Data, In-Memory Computing, Deep Learning, and GPUs. *Sustainability*, *11*(10), 2736. https://doi.org/10.3390/su11102736

Bruzzone, F., Cavallaro, F., & Nocera, S. (2023). The definition of equity in transport. *Transportation Research Procedia*, *69*, 440–447. https://doi.org/10.1016/j.trpro.2023.02.193

Chen, S., Yang, F., & Zheng, Y. (2023, October 25). *Toronto Bus Stop Optimization Using Genetic Algorithm*. Medium.

Chen, X. (2018). Review of the Transit Accessibility Concept: A Case Study of Richmond, Virginia. *Sustainability*, *10*(12), 4857. https://doi.org/10.3390/su10124857

Droj, G., Droj, L., & Badea, A.-C. (2021). GIS-Based Survey over the Public Transport Strategy: An Instrument for Economic and Sustainable Urban Traffic Planning. *ISPRS International Journal of Geo-Information*, *11*(1), 16. https://doi.org/10.3390/ijgi11010016

El-Geneidy, A., Buliung, R., Diab, E., van Lierop, D., Langlois, M., & Legrain, A. (2016). Non-stop equity: Assessing daily intersections between transit accessibility and social disparity across the Greater Toronto and Hamilton Area (GTHA). *Environment and Planning B: Planning and Design*, *43*(3), 540–560. https://doi.org/10.1177/0265813515617659

Erkul, M., Yitmen, I., & Celik, T. (2020). Dynamics of stakeholder engagement in mega transport infrastructure projects. *International Journal of Managing Projects in Business*, *13*(7).

Finio, N., Lung-Amam, W., Knaap, G.-J., Dawkins, C., & Wong, B. (2024). Equity, Opportunity, Community Engagement, and the Regional Planning Process: Data and Mapping in Five U.S. Metropolitan Areas. *Journal of Planning Education and Research*, *44*(1), 16–27. https://doi.org/10.1177/0739456X20945385

FTA. (2011). *Final Policy Statement on the Eligibility of Pedestrian and Bicycle Improvements Under Federal Transit Law*.

Goliszek, S. (2021). GIS tools and programming languages for creating models of public and private transport potential accessibility in Szczecin, Poland. *Journal of Geographical Systems*, *23*(1), 115–137. https://doi.org/10.1007/s10109-020-00337-z

Higgins, C. D., Yang, X., Widener, M., Palm, M., Vaughan, J., & Miller, E. (2022). Calculating place-based transit accessibility: Methods, tools and algorithmic dependence. *Journal of Transportation and Land Use*, *5*(1), 95–116.

Ibraeva, A., Correia, G. H. de A., Silva, C., & Antunes, A. P. (2020). Transit-oriented development: A review of research achievements and challenges. *Transportation Research Part A: Policy and Practice*, *132*, 110–130. https://doi.org/10.1016/j.tra.2019.10.018

Inwood, J. F. J., Alderman, D., & Williams, J. (2015). “Where Do We Go From Here?”: Transportation Justice and the Struggle for Equal Access. *Source: Southeastern Geographer*, *55*(4), 417–433. https://doi.org/10.2307/26233754

Jamei, E., Chan, M., Chau, H. W., Gaisie, E., & Lättman, K. (2022). Perceived Accessibility and Key Influencing Factors in Transportation. *Sustainability*, *14*(17), 10806. https://doi.org/10.3390/su141710806

Jiao, J., & Dillivan, M. (2013). Transit Deserts: The Gap between Demand and Supply. *Journal of Public Transportation*, *16*(3), 23–39. https://doi.org/10.5038/2375-0901.16.3.2

Jomehpour Chahar Aman, J., & Smith-Colin, J. (2020a). Transit Deserts: Equity analysis of public transit accessibility. *Journal of Transport Geography*, *89*, 102869. https://doi.org/10.1016/j.jtrangeo.2020.102869

Jomehpour Chahar Aman, J., & Smith-Colin, J. (2020b). Transit Deserts: Equity analysis of public transit accessibility. *Journal of Transport Geography*, *89*, 102869. https://doi.org/10.1016/j.jtrangeo.2020.102869

Kaplan, S., Popoks, D., Prato, C. G., & Ceder, A. (Avi). (2014). Using connectivity for measuring equity in transit provision. *Journal of Transport Geography*, *37*, 82–92. https://doi.org/10.1016/j.jtrangeo.2014.04.016

Kompil, M., Jacobs-Crisioni, C., Dijkstra, L., & Lavalle, C. (2019). Mapping accessibility to generic services in Europe: A market-potential based approach. *Sustainable Cities and Society*, *47*, 101372. https://doi.org/10.1016/j.scs.2018.11.047

Lin, C., Choy, K. L., Ho, G. T. S., & Ng, T. W. (2014). A Genetic Algorithm-based optimization model for supporting green transportation operations. *Expert Systems with Applications*, *41*(7), 3284–3296. https://doi.org/10.1016/j.eswa.2013.11.032

Linovski, O., Manaugh, K., & Baker, D. M. (2022). The route not taken: Equity and transparency in unfunded transit proposals. *Transport Policy*, *122*, 77–84. https://doi.org/10.1016/j.tranpol.2022.04.015

Litman, T. (2008). *Evaluating Accessibility for Transportation Planning*. www.vtpi.orgInfo@vtpi.org

Litman, T. (2018). *Evaluating transportation equity*. https://www.researchgate.net/publication/284050013

Litman, T. (2022). Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transport Planning. In *ITE Journal*. https://vtpi.org/Litman_ITEJ_Equity_Apr2022.pdf.

Liu, Y., & Cheng, T. (2020). Understanding public transit patterns with open geodemographics to facilitate public transport planning. *Transportmetrica A: Transport Science*, *16*(1), 76–103. https://doi.org/10.1080/23249935.2018.1493549

Malekzadeh, A., & Chung, E. (2020). A review of transit accessibility models: Challenges in developing transit accessibility models. *International Journal of Sustainable Transportation*, *14*(10), 733–748. https://doi.org/10.1080/15568318.2019.1625087

Manaugh, K., Badami, M. G., & El-Geneidy, A. M. (2015). Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in North America. *Transport Policy*, *37*, 167–176. https://doi.org/10.1016/j.tranpol.2014.09.013

Mesbah, M., Sarvi, M., & Currie, G. (2011). Optimization of Transit Priority in the Transportation Network Using a Genetic Algorithm. *IEEE Transactions on Intelligent Transportation Systems*, *12*(3), 908–919. https://doi.org/10.1109/TITS.2011.2144974

Murray, A. T., & Wu, X. (2003). Accessibility tradeoffs in public transit planning. *Geographical Systmes*, *5*, 93–107.

Nicoletti, L., Sirenko, M., & Verma, T. (2023). Disadvantaged communities have lower access to urban infrastructure. *Environment and Planning B: Urban Analytics and City Science*, *50*(3), 831–849. https://doi.org/10.1177/23998083221131044

Ong, P. M., Pech, C., & Green, T. (2021). *UCLA Reports Title Mobility, Accessibility and Disadvantaged Neighborhoods: Assessing Diversity in Transportation-Related Needs and Opportunities Publication Date*. https://www.metrans.org/research/mobility-accessibility-and-disadvantaged-neighborhoods-assessing-

Pereira, R. H. M., & Karner, A. (2021). *Transportation Equity*.

Sharma, S. N., Kumar, A., & Dehalwar, K. (2024). The Precursors of Transit-oriented Development. *Economic and Political Weekly*, 16–20.

Sotolongo, M. (2023). Defining environmental justice communities: Evaluating digital infrastructure in Southeastern states for Justice40 benefits allocation. *Applied Geography*, *158*, 103057. https://doi.org/10.1016/j.apgeog.2023.103057

Walls, M., Hines, S., & Ruggles, L. (2024). *Implementation of Justice40: Challenges, Opportunities, and a Status Update*.

Ward, C., & Walsh, D. (2023). “I just don’t go nowhere:” How transportation disadvantage reinforces social exclusion. *Journal of Transport Geography*, *110*, 103627. https://doi.org/10.1016/j.jtrangeo.2023.103627

Wei, R., Liu, X., Mu, Y., Wang, L., Golub, A., & Farber, S. (2017). Evaluating public transit services for operational efficiency and access equity. *Journal of Transport Geography*, *65*, 70–79. https://doi.org/10.1016/j.jtrangeo.2017.10.010

Yona, M., Birfir, G., & Kaplan, S. (2021). Data science and GIS-based system analysis of transit passenger complaints to improve operations and planning. *Transport Policy*, *101*, 133–144. https://doi.org/10.1016/j.tranpol.2020.12.009

Zannat, K. E., Adnan, M. S. G., & Dewan, A. (2020). A GIS-based approach to evaluating environmental influences on active and public transport accessibility of university students. *Journal of Urban Management*, *9*(3), 331–346. https://doi.org/10.1016/j.jum.2020.06.001

Appendix A

Main Functions of Code

```
# Main function
def main():

    batch_size = 10
    population_size = 10
    number_bus_stops = 33
    proximity_constraint = 15 / 111000  # Convert 15 meters to decimal degrees
    num_generations = 10
    mutation_rate = 0.01
    original_placement = layer_stops
    num_top_solutions = 3

    # Define the coordinate system
    coordinate_system = arcpy.SpatialReference(32613)  # Define the coordinate system (UTM Zone 13N)

    #Initial Population
    population = initialize_population(population_size, number_bus_stops, line_feature_path, proximity_constraint, coordinate_system,
original_placement)

    #Initialize GA
    best_solutions = genetic_algorithm(population, batch_size, num_generations, mutation_rate, population_size, num_top_solutions)

    # Specify a different directory
    output_path = "C:\\Users\\emili\\Documents\\best_solution.json"

    # Write best solution to a JSON file
    with open(output_path, "w") as file:
        json.dump(best_solutions, file)
        print("Best Solutions saved as a JSON file")
        print("End of Process")

```

```
# Initialize
def initialize_population(population_size, num_bus_stops, line_feature_path, proximity_constraint, coordinate_system,
original_placement):
    population = {}

    # If original placement is provided, convert it to the format expected by the population dictionary
    if original_placement:
        # Initialize an empty list to store coordinates
        original_coordinates = []

        # Open a search cursor to iterate over the point features
        with arcpy.da.SearchCursor(original_placement, ["SHAPE@XY"]) as cursor:
            for row in cursor:
                # Extract X and Y coordinates from the tuple returned by the cursor
                x, y = row[0]
                original_coordinates.append((x, y))

        # Store the coordinates in the population dictionary with the index 0
        population[0] = original_coordinates

    for i in range(1, population_size):
        # Generate a set of potential bus stop locations
        solution = generate_possible_stops(num_bus_stops, line_feature_path, proximity_constraint)
        population[i] = solution

    return population
```

```
def genetic_algorithm(population, batch_size, num_generations, mutation_rate, population_size, num_top_solutions):

    for _ in range(num_generations):

        # Dynamically set the batch size based on the number of solutions in the population
        batch_size = len(population)

        # Evaluate the fitness of the initial population
        solution_scores = fitness_evaluation(population, batch_size)  # Pass coordinate_system

        # Extract fitness scores from the solution_scores dictionary
        population_fitness_scores = {solution_id: sum(scores.values()) for solution_id, scores in solution_scores.items()}

        # Perform selection using roulette wheel selection
        selected_solutions = roulette_wheel_selection(population_fitness_scores)

        # At this point, 'selected_solutions' contains the solutions selected for the next generation
        print("Selected solutions for the next generation:")
        for i, solution_id in enumerate(selected_solutions, start=1):
            print(f"Solution {i}: ID: {solution_id}")

        # Ensure that the number of selected solutions is even
        if len(selected_solutions) % 2 != 0:
            selected_solutions = selected_solutions[:-1]  # Discard the last solution

        # Generate offspring through crossover and mutation
        offspring = generate_offspring(selected_solutions, population, mutation_rate)

        print("These are the offspring:", offspring)
        print("This is the original population:", population)

        # Combine and evaluate the original sets and offspring
        max_population_index = len(population)
        for index, solution in offspring.items():
            population[max_population_index + index] = solution  # Combine original population with offspring
        print(population)

                # Create a new set of random solutions
        random_population_size = 3
        number_bus_stops = 33
        proximity_constraint = 15 / 111000  # Convert 15 meters to decimal degrees

        # Define the coordinate system
        coordinate_system = arcpy.SpatialReference(32613)  # Define the coordinate system (UTM Zone 13N)

        # New Random Population
        original_placement = None
        new_population = initialize_population(random_population_size, number_bus_stops, line_feature_path, proximity_constraint, coordinate_system, original_placement)
        print("new_population:", new_population)

        # Add new_population to the existing population
        next_index = len(population)
        for key, value in new_population.items():
            population[next_index] = value
            next_index += 1

        new_solution_scores = evaluate_all_solutions(population)
```

```
        
        # Select only the original population size from the combined population
        new_selection = roulette_wheel_selection(new_solution_scores)[:population_size]
        
        # Create a new dictionary to store selected solutions
        population_selec = {}
        
        # Retrieve selected solutions based on new_selection
        for i, solution_id in enumerate(new_selection):
            population_selec[i] = population[solution_id]
        
        # Overwrite the original population with the selected solutions
        print("Selected solutions for the next generation:")
        print("population_selec:", population_selec)
        population = population_selec
        
        # Print population keys after adding new solutions
        print("Population keys after adding new solutions:", population.keys())
        
        print("End of cycle")
        
```

```
# Perform the evaluation on the point features
def evaluate_point_features(solution):
        maps = aprx.listMaps()
        for map_obj in maps:
            
            # Create a unique name for the output feature class
            output_feature_class = "SolutionPoints_" + time.strftime("%Y%m%d%H%M%S")
            print("Output Feature Class Name:", output_feature_class)
            
            # Create a feature class for the solution points
            solution_id_field = create_point_feature_class(arcpy.env.workspace, output_feature_class)
            
            # Insert points into the feature class for the current solution
            insert_points_into_feature_class(solution, output_feature_class, solution_id=1)
            
            # Create Buffer
            buffer_analysis(output_feature_class, output_buffer, buffer_distance)
            #dissolve_buffer(output_buffer, dissolved_buffer)
            
            # Make intersection between map and buffer
            layer = map_obj.listLayers(data_map)[0]
            clear_selection(layer)
            intersect_layers(output_buffer, layer, data_map_inter)
            
            # Create a copy of Shapefile and calculate area
            layer = map_obj.listLayers(shp_map)[0]
            clear_selection(layer)
            copy_layer(layer, shp_map_copy)
            calculate_geometry(shp_map_copy, "Area")
```

```
    # Create a copy of Intersection and calculate Buffer Area
    clear_selection(data_map_inter)
    copy_layer(data_map_inter, inter_copy)
    calculate_geometry(inter_copy, "BuffArea")

    # Sum areas and buffer areas by GEOiD to account for repetition in IDs
    statistics_fields = [["Area", "SUM"]]
    case_field = [["GEOiD"]]
    statistics(shp_map_copy, shp_sum, statistics_fields, case_field)
    statistics_fields = [["BuffArea", "SUM"]]
    case_field = [["GEOiD"]]
    statistics(inter_copy, inter_sum, statistics_fields, case_field)

    #
    fields = "SUM_Area"
    ensure_delete(data_table,fields)
    join_area(data_table, "GEOiD", shp_sum, "GEOiD", [fields])
    fields = "SUM_BuffArea"
    ensure_delete(data_table,fields)
    join_area(data_table, "GEOiD", inter_sum,"GEOiD", [fields])

    expression = "0 if !SUM_Area! == None else !SUM_Area!"
    field = "SUM_Area"
    null_cero(data_table, field, expression)

    expression = "0.0 if !SUM_Area! == 0 else float((!SUM_BuffArea! or 0)) / (!SUM_Area!)"
    field = "Coverage"
    ensure_field(data_table, field)
    calculate_ratio(data_table, field, expression)
```

```
    expression = "!P1_001N! * !Coverage!"
    field = "POP_Served"
    ensure_field(data_table, field)
    calculate_serv(data_table, field, expression)

    expression = "!P1_003N! * !Coverage!"
    field = "White_Served"
    ensure_field(data_table, field)
    calculate_serv(data_table, field, expression)

    expression = "!P1_004N! * !Coverage!"
    field = "Black_Served"
    ensure_field(data_table, field)
    calculate_serv(data_table, field, expression)

    expression = "!P1_006N! * !Coverage!"
    field = "Asian_Served"
    ensure_field(data_table, field)
    calculate_serv(data_table, field, expression)

    expression = "!Total_Labor_Force! * !Coverage!"
    field = "TLF_Served"
    ensure_field(data_table, field)
    calculate_serv(data_table, field, expression)

    expression = "!Labor_Force_Unemployed! * !Coverage!"
    field = "LFU_Served"
    ensure_field(data_table, field)
    calculate_serv(data_table, field, expression)
```

```
                expression = "!Poverty! * !Coverage!"
                field = "Poverty_Served"
                ensure_field(data_table, field)
                calculate_serv(data_table, field, expression)

                expression = "!P2_002N! * !Coverage!"
                field = "Hispanic_Served"
                ensure_field(data_table, field)
                calculate_serv(data_table, field, expression)

                clear_selection(data_map_inter)

                statistics_fields = [["P1_001N", "SUM"], ["P1_003N", "SUM"], ["P1_004N","SUM"], ["P1_006N", "SUM"], ["P2_002N", "SUM"], ["Total_Labor_Force", "SUM"], ["Labor_Force_Unemployed", "SUM"], ["Poverty", "SUM"],
                                    ["POP_Served", "SUM"], ["White_Served", "SUM"], ["Black_Served", "SUM"], ["Asian_Served","SUM"], ["Hispanic_Served", "SUM"],["TLF_Served", "SUM"], ["LFU_Served", "SUM"], ["Poverty_Served", "SUM"]]
                statistics_2(data_table, data_percent, statistics_fields)

                expression = "!SUM_POP_Served! / !SUM_P1_001N! * 100"
                field = "POP_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
                expression = "!SUM_White_Served! / !SUM_P1_003N! * 100"
                field = "White_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
                expression = "!SUM_Black_Served! / !SUM_P1_004N! * 100"
                field = "Black_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
```

```
                expression = "!SUM_Asian_Served! / !SUM_P1_006N! * 100"
                field = "Asian_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
                expression = "!SUM_Hispanic_Served! / !SUM_P2_002N! * 100"
                field = "Hispanic_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
                expression = "!SUM_TLF_Served! / !SUM_Total_Labor_Force! * 100"
                field = "TLF_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
                expression = "!SUM_LFU_Served! / !SUM_Labor_Force_Unemployed! * 100"
                field = "LFU_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
                expression = "!SUM_Poverty_Served! / !SUM_Poverty! * 100"
                field = "Poverty_Percent"
                ensure_field(data_percent, field)
                calculate_percent(data_percent, field, expression)
```

```
def roulette_wheel_selection (scores):
    total_fitness = sum(scores)
    relative_fitness = [score / total_fitness for score in scores]
    cumulative_probabilities = [sum(relative_fitness[:i+1]) for i in range(len(relative_fitness))]
    random_numbers = [random.random() for _ in range(len(scores))]

    selected_solutions = []
    for rand in random_numbers:
        for i, prob in enumerate(cumulative_probabilities):
            if rand <= prob:
                # Append the ID of the selected solution
                selected_solutions.append(i)
                break

    return selected_solutions

def one_point_crossover(parent1, parent2):
    crossover_point = random.randint(1, min(len(parent1), len(parent2)) - 1)
    child = parent1[:crossover_point] + parent2[crossover_point:]
    return child
```

```python
def generate_offspring(selected_solutions, population, mutation_rate):
    offspring = {}
    # Pair solutions for crossover and mutation
    num_pairs = len(selected_solutions) // 2
    for i in range(num_pairs):
        solution_id1 = selected_solutions[2*i]
        solution_id2 = selected_solutions[2*i + 1]

        # Convert solution IDs to integers
        solution_id1 = int(solution_id1)
        solution_id2 = int(solution_id2)

        parent1 = population[solution_id1]
        parent2 = population[solution_id2]

        # Debugging prints
        print("Parent 1:", parent1)
        print("Parent 2:", parent2)

        offspring[i] = crossover_and_mutation(parent1, parent2, mutation_rate)
    return offspring
```

Vita

Emiliano Del Rio Reyes, from Ciudad Juarez, Mexico, started his bachelor’s degree in civil engineering at The University of Texas at El Paso in 2018. During his bachelor's, he had the opportunity to explore college life by joining multiple student organizations, not only as a member but also as an officer. Despite his active engagement, he found himself uncertain about his career path, lacking a sense of fulfillment. However, his passion for research led him to join Dr. Raheem's research lab, Safe and Sustainable Construction, where he worked on projects centered around transit equity. This experience sparked his curiosity, leading him to make the decision to apply for a master’s degree in civil engineering with a focus on transportation. In December 2022, he successfully earned his Bachelor of Science in Civil Engineering with honors and secured admission to UTEP's graduate school. Emiliano has presented his research focused on transit equity at the 2023 Transportation and Construction Conference and the 2024 Institute of Transportation Engineers Conference. After graduation, he will start his professional career with a company specializing in architecture, engineering, environmental, and construction services.

Contact Information: Emiliano.del_rio@outlook.com