UNIVERSITY OF SCIENCE AND TECHNOLOGY OF HANOI

UNDERGRADUATE SCHOOL

BACHELOR THESIS

By

Trương Quốc Cường (BI11-049)

Information and Communication Technology

Title:

Optimizing the HoanKiemAir model for pedestrian areas using Optimization techniques

External Supervisor: Alexis Drogoul - ACROSS Laboratory

Internal Supervisor: Nghiêm Thị Phương - ICT Lab

Hanoi, 2023

TABLE OF CONTENT

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to the entire ACROSS laboratory team for providing me with an inspiring and enriching environment during my internship. It has been an incredible learning experience that has shaped my academic and professional journey.

I am especially grateful to my supervisor, Baptiste Lequoy, for his unwavering guidance and support throughout the project. His expertise, mentorship, and dedication have been invaluable in helping me navigate the complexities of this endeavor. I have learned immeasurable amounts under his tutelage, and his insights have greatly contributed to the success of this project.

ABSTRACT

This research focuses on the optimization of the HoanKiemAir model for pedestrian areas using advanced optimization techniques. Urban environments increasingly face challenges related to air quality and pedestrian experiences, making effective models crucial. In this study, we explore various optimization methods, including recursive and optimization algorithms, to enhance the HoanKiemAir model's performance in pedestrian zones.

The evaluation of these techniques reveals their respective strengths and weaknesses, such as simplicity, convergence speed, and decision-making transparency. Recursive algorithms offer rapid convergence but may fall into local optima, while optimization algorithms provide efficient convergence and robust solutions.

Furthermore, this research identifies key considerations, such as preventing the model from converging to unrealistic scenarios when all roads are closed. To address this issue, specific roads, like the 1A highway, and roads not directly connected to the central HoanKiemLake area are designated to remain open.

By optimizing the HoanKiemAir model for pedestrian areas, this study contributes to the development of more effective urban planning tools, promoting improved air quality and enhancing the overall pedestrian experience in urban environments.

I/ INTRODUCTION

Air pollution is a pressing global issue that poses significant threats to public health and the environment. Rapid urbanization and industrialization have led to increased emissions from various sources, including vehicular traffic, which contributes substantially to deteriorating air quality. In densely populated urban areas, such as the Hoan Kiem district in Hanoi, Vietnam, the adverse effects of air pollution are particularly evident, affecting the well-being of residents and the sustainability of the city.

Efforts to combat air pollution often involve the implementation of stringent policies and regulations to reduce emissions. One such approach is the temporary closure of certain roads, especially during periods of high pollution, to mitigate the impact of vehicular exhaust on air quality. However, determining the optimal set of roads to close requires careful consideration and analysis, as closing major roads may cause significant disruptions to traffic flow and inconvenience to commuters.

The Hoan Kiem district, known for its historical and cultural significance, is a vibrant hub of economic activities, tourism, and daily commuting. However, it also experiences high levels of air pollution due to a dense road network, heavy traffic volume, and limited green spaces. The Hoan Kiem Air Model, a comprehensive computational model for air quality analysis, provides valuable insights into the dynamics of air pollution in the district. By simulating pollutant dispersion and considering meteorological conditions, the model estimates the Air Quality Index (AQI), which serves as an essential indicator for evaluating air quality.

To date, limited research has focused on identifying the optimal set of road closures in the Hoan Kiem district to minimize the AQI while balancing traffic efficiency. Traditional approaches have primarily relied on expert knowledge and static traffic data, which may not fully capture the complex interdependencies between road closures, traffic patterns, and resulting air pollution. However, recent advancements in optimization algorithms offer promising opportunities to address this challenge more effectively.

The potential benefits of this research are multifold. While the potential benefits of this research are diverse, it is vital to acknowledge the complexities involved. Firstly, the identification of optimal road closures can indeed lead to a substantial improvement in air quality within the Hoan Kiem district. This improvement directly translates to reduced negative health impacts associated with air pollution, thus enhancing the well-being and quality of life for both residents and visitors in the area. However, it is equally important to recognize that such measures may result in a shift of air pollution burdens to nearby districts, exacerbating health risks for their inhabitants. Therefore, a comprehensive understanding of the broader implications is necessary to implement effective and sustainable measures that promote air quality improvement without unintentionally compromising the health of other communities.

Secondly, optimizing road closures can lead to a better understanding of the intricate relationship between traffic patterns and air pollution. This knowledge can inform urban planning strategies, transportation management, and the development of sustainable mobility solutions. The findings from this study may serve as a valuable reference for policymakers, urban planners, and environmental agencies in the design and implementation of effective air pollution mitigation strategies.

Lastly, by utilizing optimization techniques, this research contributes to the advancement of computational methods for urban air quality management. The proposed decision-making framework can be generalized and adapted to other urban areas facing similar challenges, providing a scalable solution to minimize air pollution while considering traffic dynamics.

II/ OBJECTIVES

The growing urbanization trend has brought with it increased concerns over rising air pollution levels in pedestrian zones. The challenge at hand is to determine optimal road closures strategically, with the primary goal of minimizing the Air Quality Index. This endeavor isn't merely about reducing traffic; it's about fostering a sustainable and healthy urban environment. Ultimately, the study may provide valuable insights and data-driven solutions that can support urban planning decision-making processes. By leveraging general optimization techniques, we aim to strike a balance between urbanization's demands and the imperative to reduce air pollution, creating a more livable and breathable cityscape for all of its residents.

The implementation varies for each algorithm, as the generalized formula requires specific encoding processes to function properly. Given the complexity of simulating the HoanKiemAir model, this presents a computationally demanding task, making fine-tuning hyperparameters within the algorithms impractical. However, previous researches have demonstrated that the generalized hyperparameters have minimal impact on the implemented algorithms; therefore, the study's reliability remains intact. Additionally, the programs have been thoughtfully designed for easy execution, facilitating interpretation and analysis.

Each algorithm provides a unique solution, representing a scenario of road closures and its corresponding simulation-based air quality index. This structure simplifies the analysis. Evaluation is based on both runtime and the quality of the final solutions to determine which algorithm best addresses the problem.

While the solutions may appear to be disparate, the results generated from scenarios of road closures yielding the best air quality index for the simulation model have furnished a valuable understanding of the relationship between urban infrastructure and air quality.

This research exemplified novel methods for a previously unexplored challenge. By employing well-established optimization techniques in conjunction with a tangible agent-based simulation, a wealth of intriguing insights pertaining to the intricate relationships among variables that collectively shape the urban fabric have been unearthed. The study has not only expanded the boundaries of knowledge about traffic accessibility but has also shed light on potential solutions and new avenues for enhancing urban planning.

III/ MATERIALS AND METHODS

1. Study Area

The study centers on investigating the GAMA simulation model of the Hoan Kiem lake region in Hanoi. This region has been a subject of concern due to its AQI. The primary objective of the study is to assess the potential impact of extending pedestrian areas on the AQI within the Hoan Kiem Lake region, while also identifying the specific areas where such extensions could potentially minimize the AQI.

The Hoan Kiem Lake pedestrian area is exclusively open for pedestrian use during the weekends, specifically from 7 pm on Friday until 12 pm on Sunday. On weekdays, the area reverts to accommodating regular traffic flows. This arrangement presents an opportunity to examine the effects of the pedestrianization strategy on air quality during designated pedestrian hours. The study's scope encompasses a thorough analysis of the road network and the associated traffic patterns in the Hoan Kiem Lake region, considering their intricate interplay with the AQI dynamics.

Hoan Kiem Lake itself holds significant historical and cultural value, spanning approximately 12 hectares and sits at the heart of Hanoi's historical center. It stands as one of Hanoi's major scenic attractions, captivating visitors with its picturesque beauty and tranquil ambiance [Vietnam Travel].

Figure 1. Phố đi bộ Hoàn Kiếm

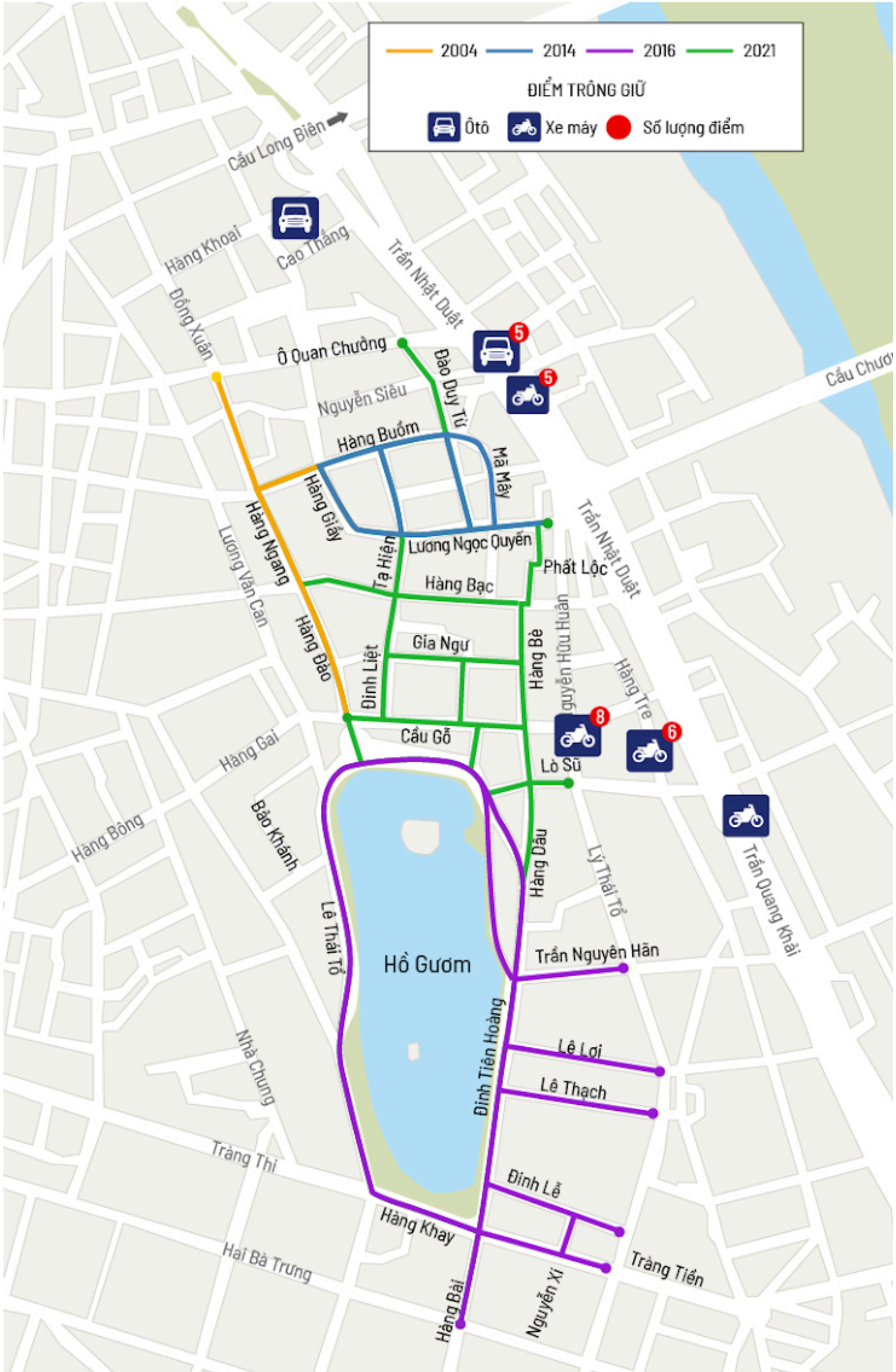

Figure 2. Current Hoan Kiem Lake pedestrian area, last expanded in 2021

By integrating the GAMA simulation platform and leveraging its capabilities in modeling and simulation, this study can shed light on the potential consequences of extending pedestrian areas within the Hoan Kiem Lake region. The findings will contribute to informing urban planning decisions and offer insights into optimizing pedestrianization strategies to mitigate air pollution and enhance the overall environmental quality of this historically significant area.

Based on a research conducted on the impact of pedestrianization on urban pollution using tangible agent-based simulations, it has been observed that closing roads for pedestrianization purposes can lead to heavy traffic congestion in surrounding areas, resulting in increased air pollution [Brugière]. However, the specific impact of extending pedestrian areas on the AQI in the Hoan Kiem Lake region needs to be examined.

The GAMA simulation platform will serve as the simulation environment for the study. By integrating the developed optimization algorithms into the GAMA simulation, the impact of extending pedestrian areas on the AQI can be evaluated effectively.

2. Materials

2.1 GAMA Platform

GAMA is a versatile open-source modeling and simulation environment that facilitates the creation of spatially explicit agent-based simulations. Its development aimed to cater to various application domains, including urban mobility, climate change adaptation, epidemiology, disaster evacuation strategy design, and urban planning. The agent-based approach advocated by GAMA promotes generality and openness, empowering users to develop customized models through plugins tailored to specific needs. Furthermore, GAMA's compatibility with other software and languages, such as R or Python, enhances its applicability, resulting in over 2000 users harnessing its capabilities for scientific simulation, scenario exploration, visualization, negotiation support, serious games, mediation, and communication tools.

The relevance and accuracy of agent-based models in GAMA are heavily reliant on the quality and accessibility of underlying data. To address this, GAMA offers a user-friendly approach to loading and manipulating Geographic Information System (GIS) data, creating an environment for artificial agents. Moreover, the platform supports the direct import and utilization of diverse data types, including CSV files, Shapefiles, OSM data, grids, images, SVG files, and 3D files (e.g., 3DS or OBJ). By enabling direct connections to databases and seamless integration with external tools like R, GAMA empowers modelers to build data-driven simulations with precision and efficiency.

While GAMA is dedicated to providing a scientific approach to model building and exploration, it remains accessible to non-computer scientists. GAML, the high-level and intuitive agent-based language within GAMA, enables users to create simulated worlds, define agent species, assign behaviors, and visualize interactions with remarkable ease and speed. A demonstration video showcasing these capabilities in under 10 minutes exemplifies GAML's user-friendly nature. Moreover, GAML caters to advanced modelers, offering a powerful agent-oriented language coded in Java. This allows for the construction of integrated models with multiple paradigms, facilitating parameter space exploration, calibration, and virtual experiments within the GAMA platform. Extensive tutorials and educational resources, coupled with ongoing support through an active mailing list since 2007, ensure that users can effectively harness the full potential of GAML, even for specialized domains like urban management, epidemiology, and risk management.

GAMA's user interface stands out as a prominent feature, streamlining the process of model creation and experiment execution. The platform allows for multiple displays within a single model, enabling users to incorporate numerous visual representations for agents and highlight crucial elements during simulations effectively. The 3D displays come equipped with support for realistic rendering, enhancing the visualization experience. GAMA further facilitates the definition of graphics for dashboard-like presentations, contributing to a more immersive simulation environment. Interactive features during simulations enable users to inspect agent populations, define user-controlled action panels, and interact with displays and external devices. The platform also offers specialized modules and plugins to enable seamless interactivity through networks, handhelds, and remote devices, enhancing the overall user experience.

GAMA's versatility and data-driven capabilities make it an ideal platform for addressing complex real-world challenges, such as determining which set of roads to close in the Hoan Kiem pedestrian area to minimize the air quality index. By leveraging GAMA's data manipulation capabilities, researchers can integrate GIS data and various datasets to create a comprehensive model of the area. The GAML language's simplicity allows for the swift definition of agent species, behaviors, and their interactions, thus constructing a realistic simulation environment. Moreover, GAMA's support for multiple modeling paradigms and parameter exploration enables researchers to test different optimization techniques effectively. By utilizing GAMA's visualization capabilities, researchers can gain insights into the impact of various road closure scenarios on air quality in the Hoan Kiem area. This approach presents a powerful tool for urban planners, environmental scientists, and policymakers to explore potential solutions to improve air quality and make informed decisions for sustainable urban development.

2.2 HoanKiemAir Model

The HoanKiemAir model provides a tangible and interactive simulation to evaluate the impacts of the pedestrian area. The pollutants emitted by vehicles are diffused on a grid and visualized on the buildings [Pham Minh Duc et al. 1].

A range of design principles is utilized to simulate traffic and pollutant emissions. It simplifies the process by assuming that vehicles only emit pollutants while in motion, with traffic being the primary source of air pollution in the simulation. This model can generate complex congestion patterns in response to an active pedestrian zone, offering a nuanced understanding of urban dynamics.

One of its key features is adaptation, allowing vehicles to dynamically adjust their routes based on real-time traffic conditions. Vehicles aim to reach their destinations as quickly as possible, taking into account factors like congestion and traffic signals. This adaptability adds a layer of realism to the simulation. Stochasticity is also introduced through the behavior of vehicle agents, making random route choices when no specific destination is set. When encountering traffic jams, vehicles can opt to change routes, providing a degree of unpredictability to the traffic flow.

The simulation incorporates detailed initialization steps, data sources for traffic and pollutant data, and sub-models for traffic, emissions, and dispersion. It calculates an AQI to simplify the interpretation of pollutant values, aiding in the assessment of air quality in the simulated area.

The AQI formula, as prescribed by governmental policies, serves as a standardized metric for quantifying air quality by taking into account various pollutant concentrations, such as particulate matters (PM), sulfur dioxide (SO_2), carbon monoxide (CO), and nitrogen oxides (NO_x) [Pham Minh Duc et al. 2]. The AQI values are systematically categorized based on the official government instructions [Vietnam Env. Admin.]. By using corresponding visualization techniques, the HoanKiemAir model provides a clear and accessible means of understanding how pollutants disperse and impact air quality within pedestrianized areas.

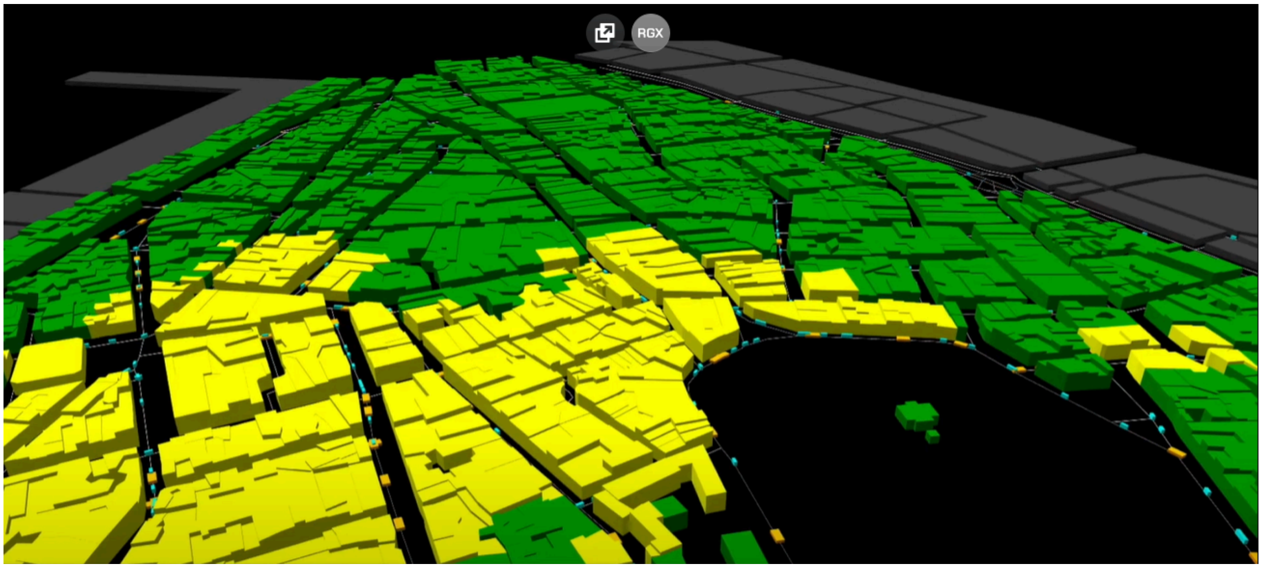

Figure 3. HoanKiemAir model from an angular point of view

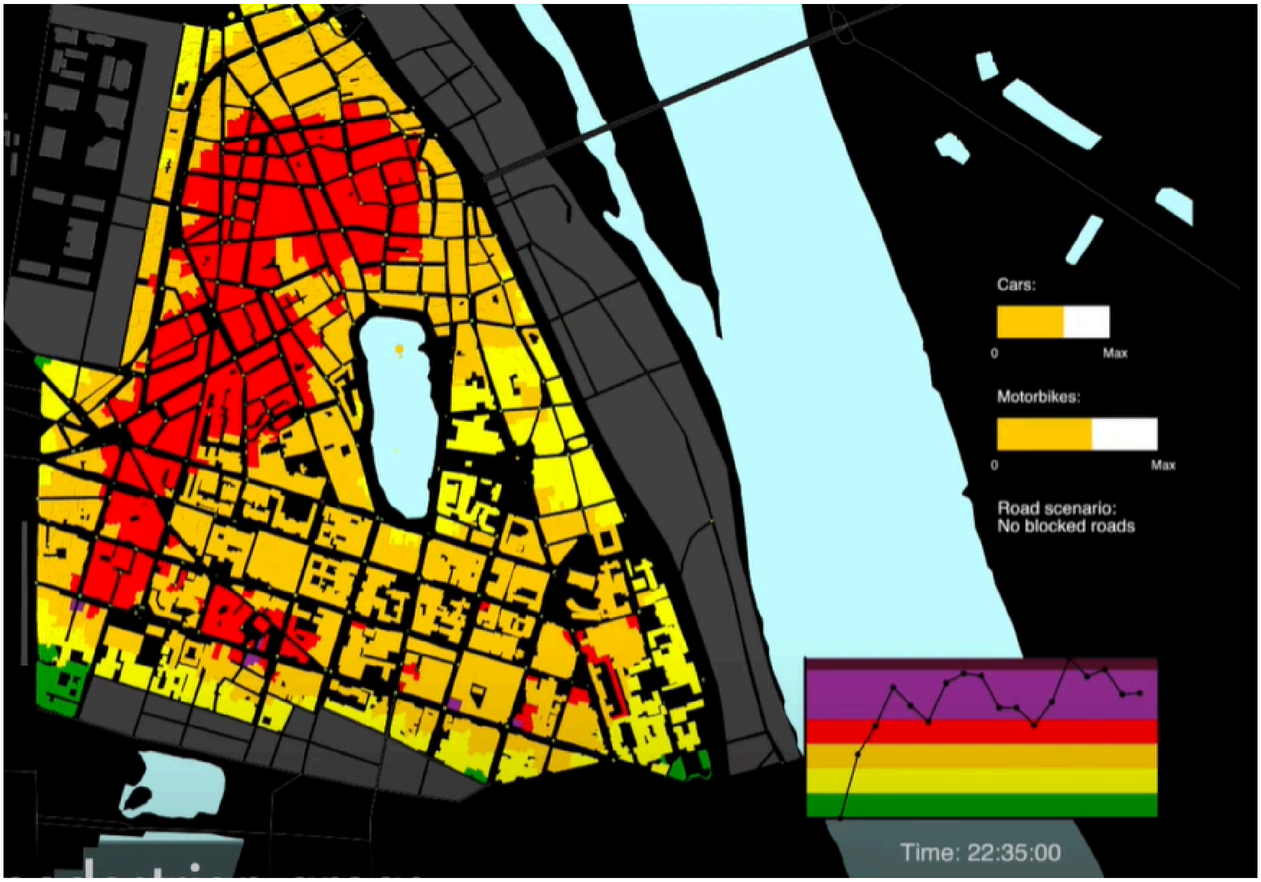

Figure 4. A simulation of the HoanKiemAir model without the pedestrian area

AQI range explanation

AQI value	Result	Description	Colors
0-50	Good	No health implications.	GREEN
51-100	Moderate	Few hypersensitive individuals should reduce outdoor exercise.	YELLOW
101-150	Polluted	Few hypersensitive individuals need reduce outdoor exercise.	ORANGE
151-200	Unhealthy	Healthy people will be noticeably affected. People with breathing or heart problems will experience reduced endurance in activities. These individuals and elders should remain indoors and restrict activities.	RED
201-300	Very unhealthy	Health effects warning: Evyerone is affected.	PURPLE
301	Hazardous	Health effects warning: Evyerone is affected.	BROWN

(*) Sensitive groups include: children, the elderly and people with respiratory disease

Figure 5. AQI range categorization

3. Methods

3.1 Implementation Process

Python serves as the programming language of choice for its ease of use and versatility.

To facilitate interaction between the developed Python algorithms and the GAMA simulation model, a straightforward web socket method is employed. This method establishes a real-time communication channel, allowing the Python programs to seamlessly communicate with the GAMA simulation environment. This approach is well-documented within the resources of the GAMA platform [GAMA Platform].

The system architecture prioritizes user-friendly automation. It empowers individuals to seamlessly execute the algorithms, oversee the GAMA simulation progress, and access the results directly in the terminal, ensuring a streamlined workflow with minimal need for manual involvement. To optimize computational resources, real-time visualization of the simulation is not employed during algorithm runtime. Instead, the simulation state is saved at the end of each simulation cycle, allowing for later retrieval and examination. This approach is valuable for error checking and results interpretation [GAMA Platform].

There is an important consideration to address before implementing the algorithms. A clear issue arises when all roads are closed, leading to an AQI of zero due to the absence of traffic in the HoanKiemAir model simulation, which might mistakenly appear as the optimal solution. To avoid this scenario, specific roads must remain open, including the 1A highway as identified in the models and roads not directly connected to the central HoanKiemLake area.

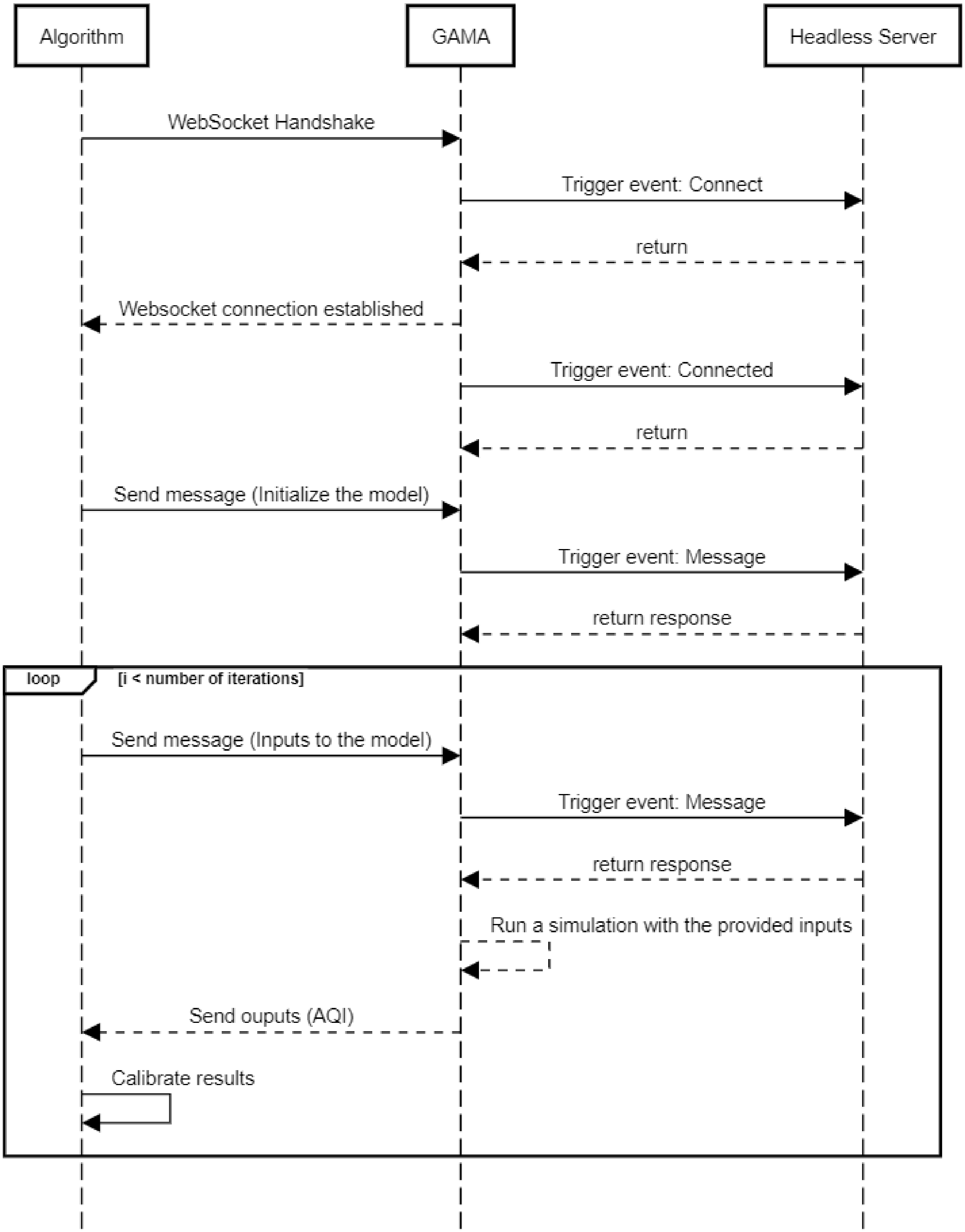

Figure 6. WebSocket Sequence Diagram

3.2 Optimization Problem Definition

The problem at hand is a combinatorial optimization problem. It involves determining the best combination of roads to close from a set of available options, with the objective of minimizing the AQI in the Hoan Kiem pedestrian area.

The action space in this problem is discrete, representing all possible combinations of road closures in the HoanKiemAir model. The Hoan Kiem district has a total of 643 roads, and 121 roads are already closed for the current established pedestrian area. Therefore, the action space would consist of 2^522 possible road closure combinations, considering the remaining 522 open roads.

3.3 Evaluation Metrics

Solution Quality, as denoted by the reduction in the Air Quality Index (AQI), stands as a pivotal indicator. The AQI Reduction metric serves as a quantifiable measure of an algorithm's ability to ameliorate air quality conditions within urban areas, thereby offering valuable insights into its effectiveness.

The AQI Reduction metric provides a robust means of gauging the tangible improvements brought about by an algorithm in terms of air quality. Its simplicity and intuitive nature make it an invaluable tool for the objective assessment of algorithmic performance. A higher AQI Reduction value signifies a more significant enhancement in air quality, providing a clear benchmark for comparing the contributions of different algorithms to environmental amelioration. This metric equips decision-makers and urban planners with a practical means of evaluating and selecting algorithms that yield the most substantial air quality improvements.

Moreover, the systematic analysis of AQI Reduction metrics plays a pivotal role in the identification of optimal algorithms for implementation in urban environments. By diligently scrutinizing the performance of various algorithms with respect to air quality enhancement, stakeholders are empowered to make data-informed decisions regarding the selection of solutions that align with their overarching objectives. In this manner, the AQI Reduction metric facilitates the prioritization of algorithms that yield the most pronounced positive effects on air quality, thereby guiding urban planning endeavors toward cleaner and healthier urban environments.

3.4 Implemented Algorithms

3.4.1 Recursive Algorithms

Within the sophisticated simulation framework of the HoanKiemAirModel, the problem of determining road closures to minimize the air quality index presents multifaceted complexities that necessitate adept decision-making strategies. One prominent approach to tackle such optimization quandaries is through the utilization of deterministic heuristic algorithms. These algorithms, built upon well-defined sets of rules or policies, offer a pragmatic way to navigate the intricate trade-offs inherent to the road closure problem.

This method of optimization ensures the algorithm addresses the most critical pollution concerns at each iteration. With a set of roads selected to be closed, the algorithm proceeds to update the road network within the simulation and recompute the air quality index. The process then continues recursively until a predetermined stopping criterion of the algorithm is met.

Although deterministic heuristic algorithms offer computational efficiency and straightforward implementation, it is crucial to recognize that their effectiveness can be constrained by the quality and suitability of the selected heuristics. As the algorithms may progressively converge to a local optimum. Consequently, their inability to identify globally optimal solutions is limited, preventing any improvements of air quality beyond the algorithms' immediate scopes. Thus, while deterministic heuristic algorithms provide valuable insights into road closure decisions, their outcomes should be interpreted with an awareness of potential limitations stemming from their deterministic and heuristic nature.

3.4.1.1 Greedy Search

The Greedy Search algorithm starts with an initial solution and evaluates its quality using a heuristic function, which, in this case, is based on the AQI of the HoanKiemAirModel. The algorithm then incrementally explores neighboring solutions by making small changes to the current solution and selecting the one that improves the heuristic the most. This process continues iteratively until a stopping criterion is met or the desired level of optimization is achieved.

Greedy Search is an attractive choice for this thesis topic due to its simplicity and ease of implementation. It provides a straightforward approach to address the problem of minimizing AQI by selecting a set of roads to close. At each step, the algorithm examines all adjacent roads, closes one road at a time, and computes the resulting AQI. The road that leads to the most significant reduction in AQI is added to the set, and the process repeats iteratively. The selection is made without considering the impact of subsequent road closures, focusing solely on the immediate improvement in AQI. This straightforward strategy allows for a quick initial exploration of the problem space, which is essential for understanding the nature of the optimization problem.

3.4.1.2 Monte Carlo Tree Search

Monte Carlo Tree Search (MCTS) represents a relatively recent addition to the repertoire of artificial intelligence algorithms, and its application in optimizing the HoanKiemAir model for pedestrian areas underscores its versatility and efficacy. Despite its apparent complexity, MCTS serves as a powerful simulation-based technique adept at seeking optimal solutions within complex, branching problem domains. This algorithm was chosen for its ability to efficiently navigate the solution space.

At its core, MCTS functions through a series of distinct steps, each contributing to the algorithm's decision-making process. The first step, known as "Selection," initiates at the root of a tree structure that mirrors the current state of the problem, in this case, the current road closure scenarios. Each of its subsequent child nodes represents an adjacent road to the current established set of roads.

The selection process hinges on a trade-off function called the Upper Confidence Bound for Trees (UCT), a pivotal concept that balances the exploration of new possibilities and the exploitation of known ones. In the context of optimizing the HoanKiemAir model, UCT evaluates the AQI of a node while considering a bias constant. This evaluation ensures efficient exploration of the solution space, effectively identifying promising paths towards air quality improvement.

Following selection, the algorithm proceeds to "Expansion." At this stage, MCTS expands the tree by considering potential actions or moves, each representing a decision that could significantly impact the problem domain. These actions encompass a spectrum of choices relevant to the optimization process, allowing the algorithm to explore a range of possibilities.

Simulation, the subsequent step, entails running multiple simulations, also referred to as playouts or rollouts, from the newly expanded nodes. These simulations involve making either random or heuristic decisions from the current state, providing estimates of potential outcomes. In the context of optimizing HoanKiemAir, these simulations mimic the impact of various decisions on air quality, enabling the algorithm to gauge the potential outcomes of different interventions.

The final step, "Backpropagation," plays a pivotal role in refining the algorithm's understanding of the quality of actions and nodes within the problem domain. The results of the simulations are systematically backpropagated through the tree, updating the statistics associated with each node. This iterative process equips MCTS with the capability to discern and prioritize high-quality actions, facilitating the search for optimal solutions.

The power of MCTS lies not only in its adaptability but also in its applicability across diverse domains. While it has found its place in board games like chess, it has also proven invaluable in tackling complex decision-making scenarios in robotics and optimization problems. Its inherent ability to efficiently explore complex solution spaces while adapting to the unique characteristics of each problem domain makes it a compelling choice for enhancing the HoanKiemAir model. By leveraging MCTS, this endeavor seeks to intelligently and systematically optimize pedestrian areas, ultimately contributing to improved air quality and urban planning decision-making.

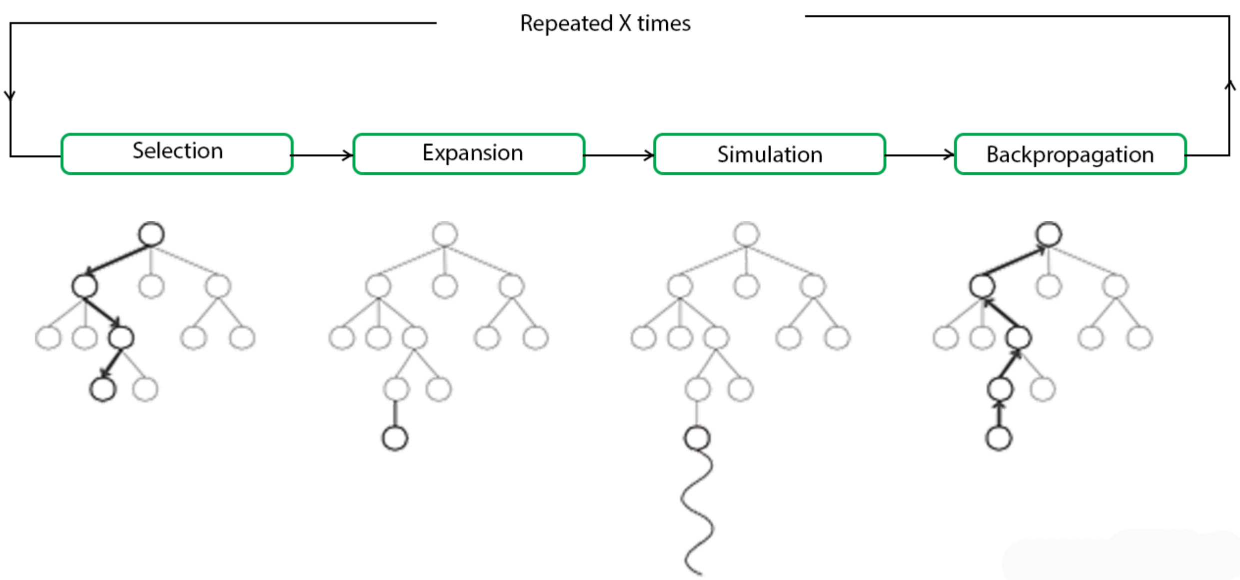

Figure 7. General Monte Carlo Tree Search process

3.4.2 Optimization Algorithms

Optimization plays a pivotal role in numerous fields, from engineering and economics to data science and urban planning. Its overarching objective is to enhance the performance, efficiency, or overall quality of systems and processes. By systematically fine-tuning variables and parameters, optimization seeks to strike a balance between competing objectives.

Conventional optimization algorithms (Deterministic algorithms) have some limitations such as single-based solutions, converging to local optima or unknown search space issues. To overcome these limitations, many scholars and researchers have developed several metaheuristics to address complex/unsolved optimization problems.

Two widely recognized optimization algorithms, Genetic Algorithms (GAs) and Particle Swarm Optimization (PSO), have been selected to solve the problem. Their applications in the study stems from their ability to efficiently search through large solution spaces, adapt to changing conditions, and provide valuable insights into the optimal configuration of pedestrian areas for improved air quality.

3.4.2.1 Genetic Algorithms

Genetic Algorithms(GAs) are metaheuristic search algorithms that belong to the larger part of evolutionary algorithms. Genetic algorithms are based on the ideas of natural selection and genetics. These are intelligent exploitation of random search provided with historical data to direct the search into the region of better performance in solution space. They are commonly used to generate high-quality solutions for optimization problems and search problems.

In the case of our study, the roads represent the possible solutions, and each combination of roads forms an individual in the population. The lists of roads indexes are encoded to binary sequences to be input in the algorithm as it represents a road being closed or not closed

The central focus is on harnessing the principles of natural selection and genetics to discover the best combination of road closures that minimizes the air quality index in the Hoan Kiem area. The algorithm mimics nature's selection mechanism, where individuals representing sets of roads with lower air pollution have a greater likelihood of reproducing and passing their road closure strategies to the next generation. The goal is to refine and evolve the combinations of roads over successive generations to identify the most effective solution for each specific scenario.

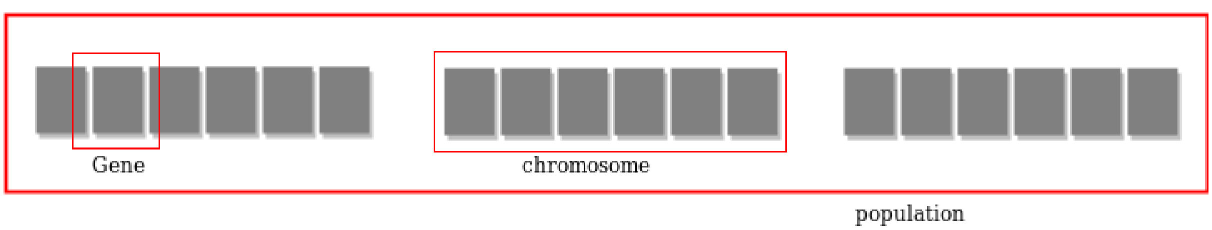

Figure 8. The search space represented in a Genetic Algorithms

Fitness Score

A Fitness Score is given to each individual which shows the ability of an individual to "compete". The individual having a low AQI on completing a simulation will have a higher fitness score. Those with highest fitness scores are sought.

The GAs maintain the population of *n* individuals (chromosome/solutions) along with their fitness scores. The individuals having better fitness scores are given more chances to reproduce than others. The individuals with better fitness scores are selected to mate and produce better offspring by combining chromosomes of parents. The population size is static so the room has to be created for new arrivals. So, some individuals die and get replaced by new arrivals, eventually creating a new generation when all the mating opportunities of the old population are exhausted. It is hoped that over successive generations better solutions will arrive while the least fit die.

Each new generation has on average more "better genes" than the individual (solution) of previous generations. Thus each new generation has better "partial solutions" than previous generations. Once the offspring produced have no significant difference from offspring produced by previous populations, the population has converged.

Operators of Genetic Algorithms

Once the initial generation is created, the algorithm evolves the generation using following operators

1. Selection Operator: This operator favors individuals (sets of closed roads) with better fitness scores, allowing them to pass their "genes" (combinations of roads) to the next generation. This mimics the concept of natural selection, where successful individuals are more likely to reproduce.

2. Crossover Operator: The crossover operator simulates mating between two individuals (sets of roads) selected through the selection operator. Random crossover sites are chosen, and their genes (specific roads) are exchanged to create entirely new individuals (offspring). This process aims to explore different combinations of roads to potentially find more effective solutions.

3. Mutation Operator: The mutation operator introduces randomness by inserting random genes (roads) into the offspring. This helps maintain diversity in the population and prevents premature convergence, ensuring that the algorithm explores a wide range of solutions.

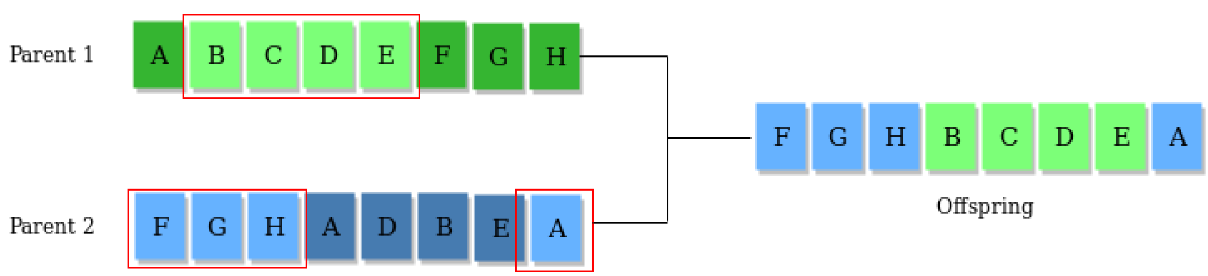

Figure 9. The mating between individuals

Elitism is applied by keeping the top 10% of the fittest individuals (lowest AQI values) from the previous generation in the new generation. These individuals are directly passed to the next generation without undergoing crossover and mutation. This ensures that the best solutions are preserved and not lost in the evolution process.

During the mating (crossover) process, each gene (road) of the offspring is determined probabilistically based on the genes of the two parents. With a probability of 45% (0.45), the gene comes from the first parent, and with another 45% (0.45) probability, the gene comes from the second parent. This accounts for a combined 90% probability that the offspring inherits genes from its parents. However, there is a 10% (0.10) probability that a gene mutates, resulting in the selection of a new random gene. This mutation operator introduces variability and helps the algorithm explore potentially better solutions that may not be present in the current population.

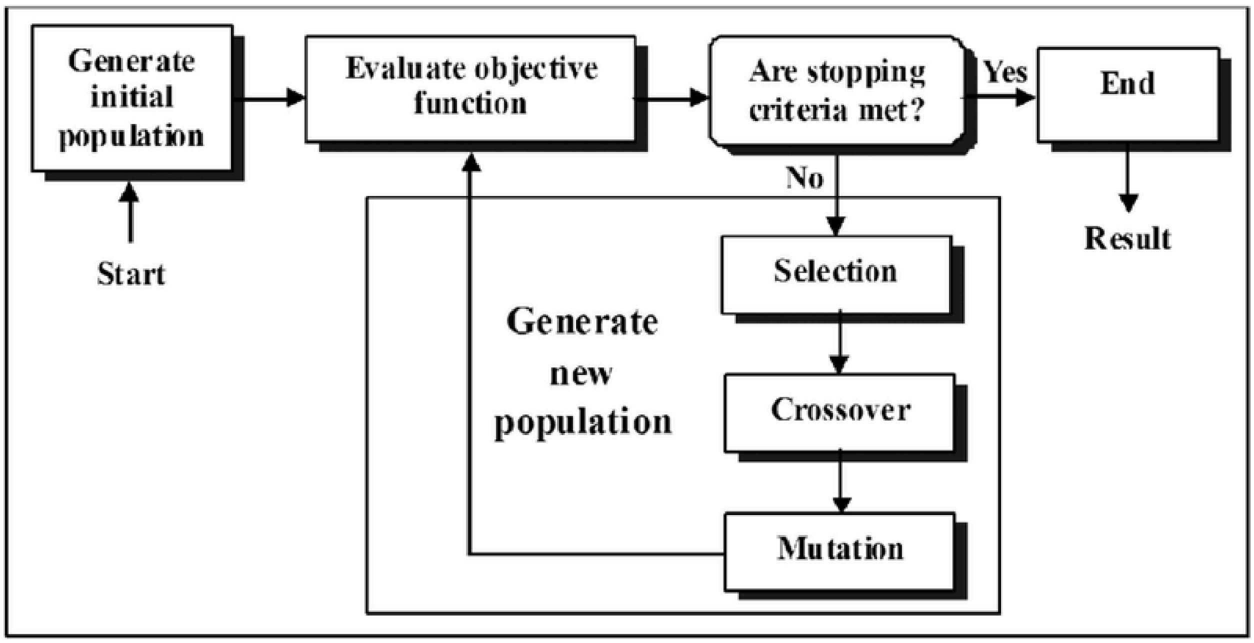

Figure 10. The general process of a Genetic Algorithms

3.4.2.2 Particle Swarm Optimization

Particle Swarm Optimization (PSO) presents itself as another popular metaheuristic algorithm of choice, distinguished by its versatility in tackling diverse optimization problems.

PSO operates iteratively, striving to enhance a candidate solution's quality based on a predefined measure of excellence. The approach revolves around a population of candidate solutions, referred to as particles, which traverse the search-space guided by straightforward mathematical equations governing their position and velocity. Each particle's journey is shaped by its knowledge of the best local position it has encountered, while also being drawn toward the most promising positions within the entire search-space. These global best positions evolve as particles discover improved solutions. The collective effect of these movements propels the swarm toward optimal solutions.

The PSO optimization process unfolds through several stages:

1. Initialization: The algorithm initializes a population of particles, with each particle's position encoding a potential solution (a set of roads to be closed) as a binary sequence. These initial positions are distributed randomly within the problem's solution space.

2. Evaluation: At this stage, each particle's position undergoes evaluation via a fitness function similar to that of the Genetic Algorithms', with a lower AQI meaning a higher fitness score.

3. Personal and Global Best Solutions: PSO keeps track of each particle's best-known position (referred to as its personal best) and disseminates information about the global best position found among all particles. These positions are updated based on the outcomes of the fitness evaluations.

4. Velocity and Position Update: Particles adjust their velocities and positions based on a mathematical formula. This formula takes into consideration their previous velocities, personal best positions, and the global best position. The purpose is to strike a balance between exploration (searching for new solution regions) and exploitation (narrowing down on promising solutions):

$$v_i = Wv_i + c_1 r_1 (P_{best,i} - x_i) + c_2 r_2 (g_{best} - x_i)$$

$$x_i = x_i + v_i$$

v_i	Velocity of the i[th] particle
x_i	Position of the i[th] particle
$P_{best,i}$	Personal best of the i[th] particle
g_{best}	Global best (Best solution of the swarm)
r_1 and r_2	Random numbers
c_1	Cognitive Acceleration Coefficient: This coefficient controls the influence of a particle's personal best information on its velocity update
c_2	Social Acceleration Coefficient: This coefficient controls the influence of the global best information on a particle's velocity update
W	Inertia weight

Figure 11. Velocity computation formula in Particle Swarm Optimization

Typically, the inertia weight w starts around 0.9 and linearly decreases to around 0.4 over the course of iterations. The common value for c1 and c2 is around 1.5. The population size chosen is in the range of 20 to 100 and the number of iterations is often set to 100 to 1000 [Clerc]. The value of the parameters can be randomized within their respective general range but have shown to have insignificant impacts on the results in many researches [Dai].

5. Termination: The PSO algorithm iteratively executes steps 2 to 4 for a defined number of iterations. This iterative process refines the particle positions and velocities, guiding the swarm toward increasingly superior solutions.

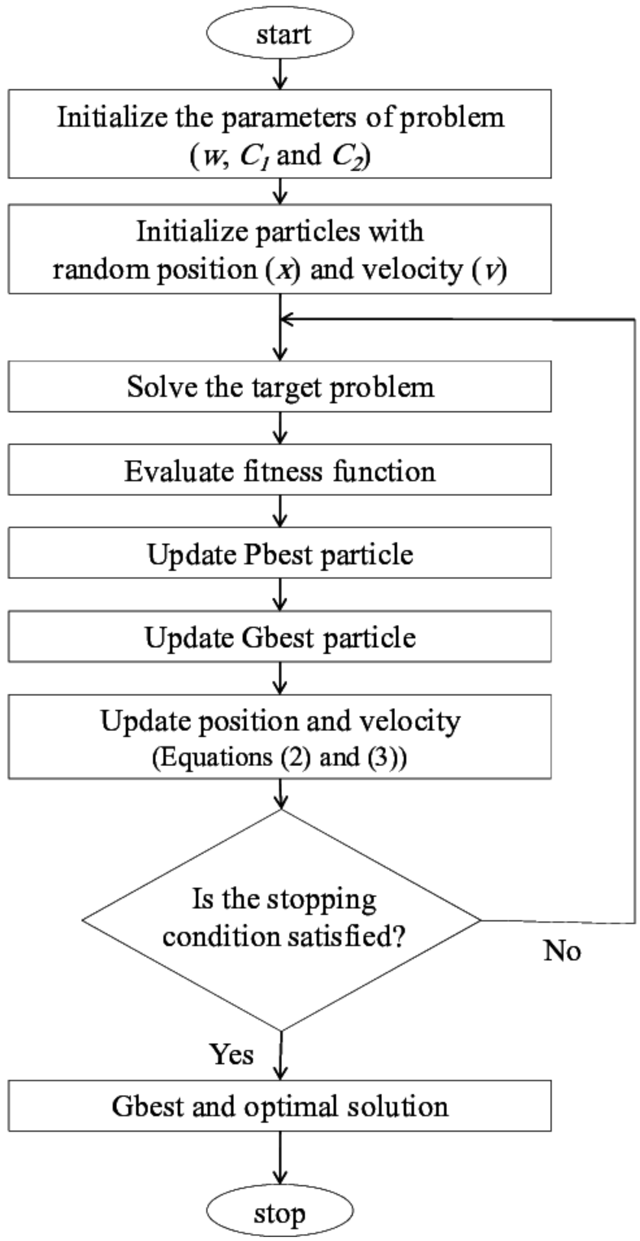

Figure 12. The general process of a Particle Swarm Optimization algorithm

IV/ RESULTS AND DISCUSSIONS

1. Results

To obtain conclusive results, three scenarios are first conducted: Scenario 1, where no roads were closed; Scenario 2, where roads within the pedestrian area were closed; and Scenario 3, where all roads except special ones were closed. These simulations spanned two days, corresponding to the weekends when the pedestrian area is accessible. To ensure the reliability of the final AQI value, 20 simulations were run for each scenario. This approach guarantees that the resulting AQI values are sufficiently consistent and not overly randomized across simulations.

Scenario	Time Ran	Number of simulations	Max AQI reached	Standard deviation of Max AQI
1. No road are closed	2 days	20	202.853	7.055
2. Roads for pedestrian area are closed	2 days	20	312.053	12.937
3. All roads beside special roads are closed	2 days	20	70.876	11.895

Table 1. AQI results from simulations of special scenarios

The runtime for each algorithm is calculated based on CPU time, with each algorithm executed 20 times. The results consider the lowest maximum AQI within their respective solutions. To ensure accuracy, the standard deviation of the results is measured, and the solution quality is evaluated for each scenario.

Algorithm	**Time Ran (CPU Time)**	**Number of times ran**	**Max AQI Reached**	**Standard deviation of Max AQI**	**Solution Quality (Max AQI Reduction) Compared to scenario 1**	**Solution Quality (Max AQI Reduction) Compared to scenario 2**	**Solution Quality (Max AQI Reduction) Compared to scenario 3**
Greedy Search	4:45:12	20	334.589	31.549	+ 131.739	+ 42.536	+ 263.712
Monte Carlo Tree Search	5:42:15	20	326.821	26.038	+ 123.971	+ 34.768	+ 255.945
Genetic Algorithms	12:53:02	20	301.580	18.967	+ 98.73	- 10.473	+ 230.703
Particle Swarm Optimization	13:13:38	20	297.510	21.509	+ 94.66	- 14.543	+ 226.634

Table 2. Solution Quality results comparison of different algorithms

To determine which roads are recommended for closure by the algorithms to reduce AQI, the lists of roads identified as optimal solutions are recorded during the algorithm runtime. Among all the generated lists of roads proposed as optimal solutions, those that appear frequently across multiple solutions may indicate the most suitable roads for closure to minimize AQI.

Percentage of occurrence in solutions	**Roads index present**
>75%	137, 90, 553, 192, 194, 575, 231, 93, 155, 239, 429, 624, 368, 44, 136, 584, 430, 286, 316, 4
50% to 75%	559, 500, 407, 472, 147, 80, 490, 89, 115, 154
25% to 50%	293, 281, 106, 247, 561, 68, 536, 503, 365, 543, 88, 615, 393, 100, 141, 345
0% to 25%	Others

Table 3. Roads that are encouraged to be closed based on solutions

2. Discussions

The analysis of Table 1 reveals compelling evidence regarding the substantial influence of pre-existing pedestrian areas on the Air Quality Index (AQI) within our simulation environment. This finding underscores the pivotal role played by pedestrian zones in shaping the air quality landscape of urban areas.

In a scenario where regular roads are strategically situated in close proximity to these designated pedestrian zones, a notably low AQI level is observed. This result can be attributed to the virtual absence of vehicular traffic within the simulation, as these specialized pathways effectively limit and redirect traffic away from the immediate vicinity. Consequently, emissions from vehicles are minimized, contributing to a favorable air quality environment.

Moreover, an examination of the standard deviation across various simulation runs reveals a noteworthy pattern. The relatively limited variance in the final AQI values among these simulations suggests a high degree of consistency in our results. This consistency is a key indicator of the reliability of our model and its ability to produce dependable AQI predictions across different scenarios.

In Table 2, several intriguing insights come to light. Firstly, a notable feature worth highlighting is the relatively high standard deviation associated with the results obtained from the recursive algorithms. This observation implies that these algorithms frequently find themselves trapped in local optima during their optimization processes. Consequently, this recurrent entrapment in local optima significantly undermines the reliability of the outcomes produced by these algorithms. Such unreliability in the results complicates their interpretation and practical application.

Furthermore, it is instructive to note that the recursive algorithms' worst-performing instances align with the hypothesis that their primary focus on immediate AQI reduction may, in fact, hinder their ability to achieve optimal solutions for the broader problem at hand. This limitation becomes evident in the recorded results, which indicate a failure to effectively reduce air quality and discover an optimal solution. This outcome is entirely consistent with the characteristics of algorithms constrained by the challenges of local optima.

Secondly, the optimization algorithms featured in the table exhibit a lengthy runtime and a slow convergence rate. These attributes suggest that these algorithms demand substantial computational resources and time to reach a solution. However, and perhaps more significantly, the AQI reduction achieved by these optimization algorithms is commendable. Despite the extended execution times, they effectively contribute to improving air quality. The standard deviations of the solutions does appear less than ideal. These standard deviations indicate variability and dispersion in the outcomes generated by the algorithms. While lower standard deviations are generally preferred as they signify more consistent results, in this context, the slightly higher standard deviations should not be dismissed outright. Instead, they shed light on the complexity of the problem at hand.

Analyzing the data provided in Table 3 reveals a set of roads that emerge as optimal candidates for closure within the model to mitigate AQI levels. These identified roads exhibit distinct characteristics; they are scattered throughout the network and lack a clear pattern of proximity or adjacency to one another. While these roads individually demonstrate efficacy in reducing AQI, the distribution and arrangement of these optimal closures appear somewhat arbitrary.

This particular distribution of optimal road closures may be considered less intuitive from a practical standpoint. Ideally, road closures intended for AQI reduction would exhibit a more logical and coherent pattern, such as targeting roads in close proximity to one another or along specific routes known to contribute significantly to pollution levels. Therefore, the current distribution of optimal closures, albeit effective individually, raises questions about the overall efficiency and cohesiveness of the road closure strategy in achieving a comprehensive reduction in AQI levels.

V/ CONCLUSION

In conclusion, this thesis has delved into the optimization of the HoanKiemAir model for pedestrian areas, employing both recursive and optimization algorithms. The comparative analysis has illuminated the strengths and weaknesses of each approach.

Recursive algorithms, characterized by their simplicity and quick convergence, offer an easily interpretable decision-making process. However, they are susceptible to getting trapped in local optima, overlooking non-immediate solutions, and can rapidly converge to suboptimal solutions.

On the other hand, optimization algorithms exhibit fairly simple implementation and efficient convergence towards near-optimal solutions. They strike a balance between exploration and exploitation, providing robust solutions. Nevertheless, they tend to be slow to converge and often involve challenges in parameter tuning.

This research has opened up avenues for further development in the optimization of the HoanKiemAir model for pedestrian areas. Two prominent directions for future exploration include the integration of parallel optimization algorithms and the incorporation of machine learning techniques, such as neural networks and reinforcement learning. Additionally, leveraging domain-specific knowledge about traffic patterns, air quality dynamics, and local conditions can lead to more precise and context-aware optimizations in pedestrian areas.

REFERENCE

[Brugière] Brugière, A., Pham, M., Chapuis, K., Drogoul, A., Gaudou, B., Grignard, A., Marilleau, N., Nguyen-Huu Tri. (n.d.). Experimenting the Impact of Pedestrianisation on Urban Pollution Using Tangible Agent-Based Simulations: Application to Hoan Kiem District, Hanoi, Vietnam. Retrieved from https://link.springer.com/chapter/10.1007/978-981-16-2629-6_4

[Clerc] Clerc, M., & Kennedy, J. (2002). The Particle Swarm - Explosion, Stability, and Convergence in a Multidimensional Complex Space. IEEE Transactions on Evolutionary Computation, 6(1), 58-73.

[Dai] Dai, H.-P.; Chen, D.-D.; Zheng, Z.-S. (2018). Effects of Random Values for Particle Swarm Optimization Algorithm. Algorithms, 11, 23. doi: 10.3390/a11020023.

[GAMA Platform] GAMA Platform. (n.d.). Headless Tutorial. Retrieved from https://gama-platform.org/wiki/HeadlessTutorial#use-the-python-wrapper-instead

[GAMA Platform] GAMA Platform. (n.d.). Save and Restore Simulations. Retrieved from https://gama-platform.org/wiki/Save-and-restore-simulations

[Pham Minh Duc et al. 1] Pham Minh Duc, Kevin Chapuis, Alexis Drogoul, Benoit Gaudou, Arnaud Grignard, et al. (2020). HoanKiemAir: simulating impacts of urban management practices on traffic and air pollution using a tangible agent-based model. In International Conference on Computing and Communication Technologies (RIVF 2020), Oct 2020, Ho Chi Minh, Vietnam, pp. 1-7. doi: 10.1109/RIVF48685.2020.9140787. [hal-02968254]

[Pham Minh Duc et al. 2] Pham Minh Duc, Kevin Chapuis, Alexis Drogoul, Benoit Gaudou, Arnaud Grignard, et al. (2020). HoanKiemAir: simulating impacts of urban management practices on traffic and air pollution using a tangible agent-based model. In International Conference on Computing and Communication Technologies (RIVF 2020), Oct 2020, Ho Chi Minh, Vietnam, pp. 1-7. doi: 10.1109/RIVF48685.2020.9140787. [hal-02968254]

[Vietnam Env. Admin.] Vietnam Environmental Administration. (2011). Quyet dinh 878/QĐ-TCMT ve viec ban hanh so tay huong dan tinh toan chi so chat luong khong khi (AQI).

[Vietnam Travel] Vietnam Travel. (n.d.). Hoàn Kiếm Lake - The Heart of Hanoi, Vietnam. Retrieved from https://vietnamtravel.com/overview-of-hoan-kiem-lake-hanoi/