Personalized Experience-aware Multi-criteria Route Selection for Smart Mobility

Matheus Brito[1], **Bruno Martins**[1], **Camilo Santos**[1], **Iago Medeiros**[1], **Felipe Araújo**[1], **Marcos Seruffo**[1], **Helder Oliveira**[2], **Eduardo Cerqueira**[1], **Denis Rosário**[1]

[1]Federal University of Pará (UFPA), Belém – PA – Brazil

[2]Federal University of ABC (UFABC), Santo André – SP – Brazil

{matheus.moraes.brito,bruno.martins,iago.medeiros}@itec.ufpa.br,

camilo.santos@icen.ufpa.br,helder.oliveira@ufabc.edu,

{felipe.araujo,seruffo,denis, cerqueira}@ufpa.br,

Abstract. *Smart urban mobility emerged from the urban citizen's need for a fast urbanization environment, using personal devices and city infrastructure integration, data generation, and mobility services provided on congested and possibly dangerous urban roads. However, traditional routing services need to consider users' experience, comfort and health because they usually choose only routes with the shortest paths or less traffic. This work proposes a route selection method based on a personalized preference for different user profiles, and essential geolocated factors from data collection, including crime occurrences and air quality factors. The suggestion method allows safer, healthier, and more pleasant paths for drivers and analytic data for city planners compared to single-criteria route selection approaches.*

1. Introduction

The extensive urbanization process and consumer demand in many world scenarios cause a delay in urban mobility evolution and represent significant global challenges. Passenger and goods transportation services imply fleet growth, causing traffic congestion and poor air quality in city areas. However, urban mobility flow improvement relies on economic-ecological sustainability allied to drivers' satisfaction attainment [Savithramma et al. 2022].

A personalized experience-aware route selection scheme must consider multiple factors in smart city transportation. For instance, technology, policy, community, and environment sectors must be consulted and attended to improve city transportation and achieve the smart transportation systems status [Legaspi et al. 2020]. Through information technologies implementation, the exchange and processing of sectors' data contribute to mobility solutions addressing all factors that affect the quality of life and business when implementing city mechanisms.

The modernization and popularity of connected devices with usage in different social areas bring ubiquity to technology. Many web-connected things interact among them and users, resulting in extensive data acquisition for a route selection scheme. In addition, the Internet of Things (IoT) represents systems and physical objects interconnected to the Internet for data exchange between heterogeneous devices without human intervention [Nguyen et al. 2022]. For instance, IoT applied to urban mobility can present smart elements, such as traffic lights, parking lots, and flow management, which provide a vast amount of data for a personalized experience-aware route selection scheme.

In this context, Geographic Information Systems (GIS) evolved with technological advances, presenting a new representation of spatial data beyond typical maps. Vehicular navigation systems apply geo-spatial and geo-referred data, which offer a dense data quantity and variety [Bellini et al. 2022]. GIS provides assistance to urban drivers, but it the lack of criminality and accident data that may lead to a potentially dangerous path.

Moreover, air pollution constitutes a more significant threat to public health, causing one in eight deaths worldwide, as 92% of the world's population lives where pollution exceeds safe limits [Brito et al. 2022]. Criminality also affects the urban reality, risking social stability and economic progress, which results in the urgent need to prevent it [Han et al. 2020]. In this context, government administration worldwide with open data policies provides access to criminality and road accident occurrence for research purposes, developing solutions for minimizing events [Liu et al. 2019]. Therefore, custom navigation solutions can help different drivers' preferences and needs, offering securer, faster, or healthier urban trips.

This paper applies the Analytic Hierarchy Process (AHP) as a decision support method for route alternative ranking and considers four personalized drivers' profile preferences, named as, Worker, Green, Safe, and Tourist, providing adaptive weights for each features. For instance, relative weights attributed to health, security, comfort, and well-being affect alternative route ranking according to a user profile. A London urban open dataset provides an Origin-Destination (OD) pair with route alternatives and normalized factors weights [Rodrigues et al. 2021]. We present a comparison between different user profile weights and "greedy" options, which consider a single criterion as the higher weight for route selection, proving the efficiency of customized preference. The evaluation shows superior statistical performance for our personalized profile method compared to greedy profiles for route selection considering all values.

This paper's organization follows: Section 2 summarizes the main related works to this paper's scope. Section 3 presents the scenario overview of the proposed methodology with decision support routing ranking, adding pollution criterion weight for selection through implementation. Section 4 presents the evaluation method for the personalized profiles ranking after implementation. Finally, Section 5 presents the paper's conclusions and future works.

2. Related Works

This section presents the main state-of-art works which approach the multi-criteria method, navigation systems, and trip influencer factors. For urban-logistic routing recommendation, [Wu et al. 2022] proposed a vehicle-route optimization approach using contextual traffic data and multi-criteria decision analysis. The author built the contextual data from the urban transportation database and Google Maps routing API metadata. The work presented an urban route selection for deliverymen used in urban logistics, needing to be more beneficial for citizen use. Besides, the criteria for selection are limited to average speed, congestion degree, distance, and worker personal interest.

[Sarraf and McGuire 2020] added an analytic choice method for the developed safer route planning application, comparing different multi-criteria methods outcomes for the same objective. Regardless, the proposed system analyzed both historical and live monitoring, considering vehicle accidents in the analysis area, offering safer routes and urban infrastructure reports for future enhancement.

[Kaivonen and Ngai 2020] proposed real-time monitoring with data gathering on pollution through urban public transportation networks, covering the whole city area, addressing air quality issues as urban environment criteria. The authors evaluated data collection on mobile sensors compared to stationary air sensors, choosing an efficient way to map pollution in the urban environment. Although, the solution does not return a less polluted route alternative for user selection.

[Zhang et al. 2022] proposed a routing method for Vehicular *Ad Hoc* Network based on the relative speed, the angle between node and neighbors, the connection angle between destination node and its neighbors, and the node density of neighbors, as criteria and combined all these criteria into a node location algorithm. Their proposal is an efficient routing approach for VANETs but did not consider contextual data for humanized mobility and only improves communication metrics between devices

[Hsieh and Lin 2022] proposed a route recommendation method for taxi drivers that considers real-time predictions and traffic network information, aiming for higher profit. The criteria for this approach rely on pick-up probability, drop-off distribution, road network, distance, and time factors. The authors did not consider health, comfort, and risk factors, only the usual navigation criteria compared to standard methods.

[Solé et al. 2022] proposed a method for feature measurement that affects driver security and pleasure on urban trips. This work evaluated different route selection methods with single or multi-criteria and some pre-defined profiles. The authors intended the best route identification from a dataset about the City of London containing weights of different trip influencing factors, but did not consider air pollution as a criterion.

The literature review indicates the integration need for other factors in vehicle trip suggestion, using emerging technology to enhance the data acquisition step for route selection from each driver's necessities. Table 1 shows the relation between previous works and this paper on different issues, such as the multi-criteria approach, various criteria in the selection, including air pollution, and providing the best route ranking based on defined user profiles preference. This paper presents contributions on each element integration, contrasting state-of-art approaches.

Table 1. Features addressed in related works

Work	Multi-criteria Approach	Air pollution factor	Comfort and security factors	Routes ranking	User Profiles
[Wu et al. 2022]	✓	✗	✗	✓	✗
[Sarraf and McGuire 2020]	✓	✗	✓	✓	✗
[Kaivonen and Ngai 2020]	✗	✓	✗	✗	✗
[Zhang et al. 2022]	✓	✗	✗	✗	✗
[Hsieh and Lin 2022]	✓	✗	✗	✓	✗
[Solé et al. 2022]	✓	✗	✓	✓	✓
Proposed Methodology	✓	✓	✓	✓	✓

3. Multi-criteria Route Selection

This section describes how our method works, combining the multi-criteria decision-making method and the urban routes' contextual data. We also define the necessary steps for validating the AHP criteria preferences and introduce the planned user profiles. At least, we detail the method application specifying the implemented algorithm, highlighting its efficiency.

3.1. Scenario Overview

Figure 1 presents the overview for a personalized experience-aware multi-criteria route selection scheme. The criteria definition step processes all contextual route-related data. The selection method step defines the AHP method criteria and alternatives validation with user profile weights description. Finally, the method evaluation steps evaluate the method application by comparing the custom profiles with greedy profiles.

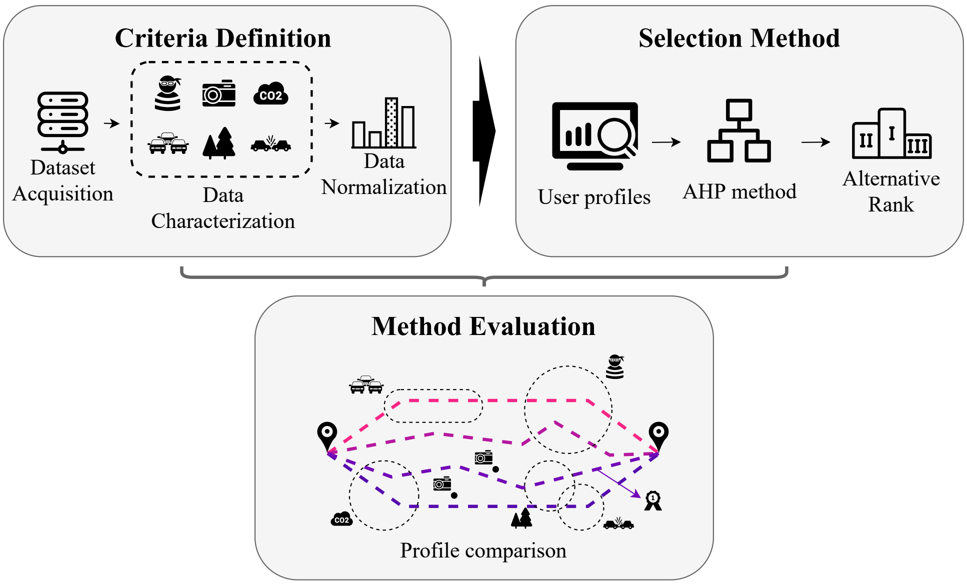

Figure 1. Methodology overview for route selection

The data acquisition phase in the criteria definition step consists in retrieving all contextual and physical data for the dataset build. The authors [Rodrigues et al. 2021] collected open data from websites and geographic tools to complete the dataset. We added the pollution factor through the local air quality open database [Kelly and Kelly 2009]. In the Data characterization phase, we insert all contextual and physical feature values to the route alternatives, updating The London routes dataset with the pollution level. The updated dataset contains eight criteria elements, described as follows:

- **Crime**: This criterion is related to the criminality level considering crime event history in determined areas. An open data United Kingdom police repository [Bibri and Krogstie 2020] containing all geolocated crime records are analyzed, and the average crime severity assigns the crime criterion value.
- **Accidents**: Defines a danger level to vehicle accidents near a determined route. The United Kingdom government's open data repository [Bibri and Krogstie 2020] provides geolocated accident records. The accident severity degree and the fatalities that occurred define the criterion value.

- **Nature**: Natural landscapes and "green" areas affect trip aesthetics. Parks, gardens, marinas, golf fields, nature reserves, lawns, meadows, and water define a pleasant trip and decrease driver stress. The Overpass API provides the natural occurrence through OpenStreetMap API [OpenStreetMap 2017], allowing the criterion value through the intersected area between nature polygons.
- **Attractions**: Defines the tourist attractions near the route traces. Overpass API [OpenStreetMap 2017] provides geolocated Points-Of-Interest (POI) data. The attraction level indicates the POIs number in the region.
- **Duration**: Defines a traditional parameter for a vehicular navigation system affecting driver trip perception. Long trips may be a stressful experience and widely avoided. HERE API [HERE 2023] provides the estimated duration for each route for alternative route tuple adding.
- **Traffic**: Represents the most stress-related trip factor, implying in-route travel time. HERE API [HERE 2023] provides the route vehicle density level. The traffic level is the comparison between duration with and without vehicle density.
- **Length**: Navigation system elementary factor provided by HERE Maps API [HERE 2023], directly impacting the internal combustion engine vehicle consumption and travel financial cost.
- **Pollution**: We added the pollution factor for the dataset with the London Air Quality Open Data [Kelly and Kelly 2009]. The raw values the NO_2 concentration level near the route, with a $300m$ sensor tolerance, which affects users' health.

The data normalization phase standardizes each criteria raw value from 0 to 1 for multi-criteria application on alternative selections. The Equation 1 calculates the normalized value for each criterion raw value (X_i) in the dataset, with X_{max} representing the maximum value and X_{min} the minimum value, which the lower occurrence indicates a better index for selection, such as crime and accidents occurrence, estimate duration, trip length, and pollution level. The Equation 2 normalizes the criterion raw value inversely for some criteria for which the higher occurrence indicates a better index for selection, such as natural areas, tourist attractions, and traffic ratio.

$$\widehat{X_i} = \frac{X_{max} - X_i}{X_{max} - X_{min}} \tag{1}$$

$$\widehat{X_i} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{2}$$

In the user profile phase of the selection method step, we define the pairwise preference comparison between features for the four profiles (Worker, Green, Safe, and Tourist). Afterward, the AHP method phase guarantees the matrix consistency for preference if it needs any correction. The alternative rank phase defines the best route selection as the product between the preference and alternative weights.

Finally, the method evaluation step analyses the best result for all routes for the user and greedy profiles under a profile comparison, corresponding to selection preference with higher priority on only one feature. These comparison objectives validate the user preference as a practical way to make a route choice.

3.2. Selection Method

For the correct route selection among alternatives with different contextual information, we need to choose a decision-making method that is simple and robust, aiming at scalability for any urban environment application. The AHP method for route selection offers robustness, considering dense criteria and weight with simple mathematical calculations for selection. Figure 2 shows the route selection objective hierarchy, complementing traditional car navigation system factors with pleasant and health factors. As initially arranged in the dataset, alternative paths for OD pair sets range from two to seven selection paths.

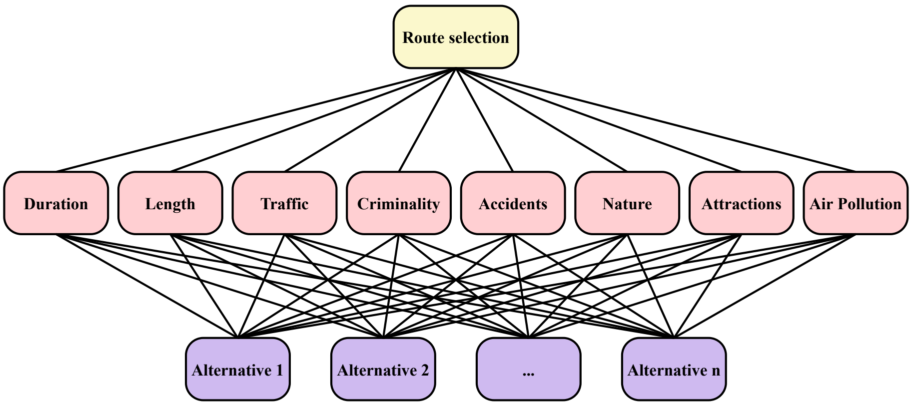

Figure 2. Hierarchy model for route selection representation

The built hierarchy between elements and the weight definition to each alternative route tuple defines the path preference order. [Saaty 1990] describes the correct AHP use with decision matrix determination for every alternative relating to a criterion, defining the criteria normalized indexes. Nonetheless, the collected dataset for analysis normalizes its raw value indexes for each criterion, allowing the full use of the AHP method.

Saaty's scale [Saaty 1990] defines the element importance degree to another and allows the comparison matrix built, as shown in Equation 3. M represent the decision matrix with all $f_{n,n}$ pairwise comparison. The matrix objectives are problem complexity level reduction and driver's profile preference definition, facilitating method application due to elevated criteria quantity in various problems.

In the pairwise comparison, the AHP method uses a verbal judgments scale ranging from "equal" to "extreme" (equal relevance, great relevance, greater relevance, huge relevance, and extreme relevance), referring to a criterion comparison importance to another for the problem solution reach. Numerical judgments represent every verbal judgment, being "equal" equivalent to 1 and "extreme" to 9 (1, 3, 5, 7, and 9) with the intermediate values (2, 4, 6, and 8).

$$M = (F_{i,j})_{n \times n} = \begin{pmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,n} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n,1} & f_{n,2} & \cdots & f_{n,n} \end{pmatrix} \tag{3}$$

A consistency validation for the matrix is an AHP method step for the correct matrix built. Equation 4 shows consistency ratio (CR) and consistency index (CI) fraction to obtain random index (RI), calculating decision-makers' judgments consistency.

$$CR = \frac{CI}{RI} \tag{4}$$

The maximum matrix eigenvalue (λ_{max}) must be equal to matrix dimension *n* for matrix consistency maintenance. The $n - 1$ value is used for logically deduced pairwise comparison. Therefore, the fraction between these elements obtains the Consistency Index (CI), shown in Equation 5.

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{5}$$

The maximum eigenvalue indicates the judgment consistency measure, calculated through the judgment matrix (A) and the priority column vector (*w*) product, which splits the vector mean, as seen in Equation 6.

$$\lambda_{max} = \text{vector mean} \frac{Aw}{w} \tag{6}$$

The author also defines the random index as a constant value applied to defined decision matrices for the hierarchy analysis method. This paper uses the 1,41 random index value for using an eight elements matrix [Saaty 1990]. The formulas calculation obtained valid consistency index for all profiles since it is less than 10%.

We define the four user preferences for the method application: Worker, Green, Safe, and Tourist. In this way, we define four profile matrices to achieve the relative weights for further method application, as shown in Table 2. The result weights for Higher criteria weights indicate a higher preference, while smaller criteria indicate the opposite. The alternative evaluation process will use criteria weights for selection.

Table 2. Trip feature weights for different profiles

AHP Profile	Crime	Accident	Nature	Attraction	Duration	Traffic	Length	Pollution
Worker	0.046	0.063	0.021	0.021	0.260	0.227	**0.328**	0.034
Green	0.085	0.040	0.280	0.087	0.031	0.055	0.059	**0.362**
Safe	**0.369**	0.244	0.024	0.023	0.073	0.129	0.047	0.092
Tourist	0.164	0.101	0.117	**0.394**	0.058	0.018	0.011	0.044

For instance, the Worker profile has a higher weight in the Length feature, followed by Duration and Traffic, aiming for faster trips. The Green profile feature rank is Pollution and Nature for a bucolic and healthier trip. Safe profile seeks a securer trip, prioritizing Crime occurrence and Accidents. At least, the Tourist profile is for travelers and visitors, with higher weights on the Attraction feature. For the research purpose, we consider only four profiles for demonstration; the method can work with any preference as long as the matrix is valid for the AHP method.

3.3. Method application

In this section, we execute the method application through the AHP consistency measure and the ranking acquire. Furthermore, we use all the acquired results to calculate the average best option for each user profile.

Algorithm 1 computes the method application for matrix consistency calculation and the route evaluation for each OD pairs in order. The interval `line 1 - line 4` declares the consistency index, the consistency ratio, and the F value for matrix weights . In `lines 5` to `14`, we applied the AHP method for attributing the preference weights. In `lines 15` to `17`, we define the variables for alternative evaluation. At least, in `lines 18` to `20`, we calculate the alternative route performance by multiplying every preference tuple value with each normalized feature value, returning the result array for further comparison, as seen in `line 21`.

Algorithm 1 Decision matrix consistency and route evaluation

Require: M,C
1: $incRat \leftarrow 1.41$
2: $consistencyRatio \leftarrow 0.10$
3: $F \leftarrow shape(M, 1)$
4: $weights \leftarrow$ a list of zeros in the range of F
5: **for** $i \leftarrow 1$ to $length(F)$ **do**
6: $\quad weights[i] \leftarrow reduce(F(x, y) = x.y, M[i, :]^{(1/F)})$
7: **end for**
8: $weights \leftarrow weights/$sum of all elements in weights
9: $\lambda_{Max} \leftarrow mean(sum(M.weights)/weights)$
10: $consInd \leftarrow (\lambda_{Max} - F)/(F - 1)$
11: $RC \leftarrow consInd/incRat$
12: **if** $RC > consistencyRatio$ **then**
13: $\quad$ **return** $\emptyset$
14: **end if**
15: $routesParams \leftarrow C[columns[parameters]]$
16: $resultsArray \leftarrow \emptyset$
17: $N \leftarrow length(routesParams)$
18: **for** $i \leftarrow 1$ to N **do**
19: $\quad resultsArray \leftarrow sum(multiply(routesParams[i], weights))$
20: **end for**
21: **return** $resultsArray$

For computational method implementation and flexible route selection, many programming tools achieve geolocated data filtering and route alternative order definition goals. The method framework aims at data analysis of factors, including pollution, and inserting each alternative tuple. Each criteria-defined value distinguishes the best route and the alternative order for any OD pair.

The AHP method application, jointly with the route evaluation algorithm, as shown in the Algorithm 1, presents a time complexity, in the worst case, as O*(n*m)*, where *m* is the number of features to be evaluated in the profile, and *n* represents the number of alternative routes within an OD pair. We consider the presented complexity efficient due to its asymptotic value being limited by a polynomial.

4. Evaluation

In this section, we describe the dataset and introduce the mechanism of pollution level attribution for paths near air quality sensors. Furthermore, we define the statistical tool for profile comparison to establish our personalized approach efficiency.

4.1. Methodology

The London routes were designed for selection methods evaluation, containing different factors besides the standard time, length, and traffic. The criminality, accidents, nature, and attractions metrics consideration imply more pleasant and safe trips through the city. To consider the drivers' health and well-being, we introduce the air quality attribution to routes through sensor readings and add the pollution value to the dataset.

In this way, we consider the London public pollution data, which allows air quality level attribution for each alternative route with collected readings timestamp in the same dataset date. London Air Quality Network (LAQN) API provides pollution sensor readings, with sensors installed in and around London. Integrated sensors network has a real-time data collection of main pollution-related gaseous substances: ozone (O_3), nitrogen dioxide (NO_2), and inhalable particles with a diameter smaller than 10 and 2,5 micrometers. The API request retrieves the 2020 readings information, with a significant presence of NO_2. Pollution feature considers NO_2 level, with normalization for route selection method application.

We assume a set of routes in the dataset as a latitude-longitude pair path set $R = \{1, ..., n\}$, $R \in \mathbb{R}^{n \times 2}$. The methodology considers $C = \{R | R \in \mathbb{R}^{n \times 2}\}$ a path set with standard departure and arrival OD pair, and then it is considered $R_i \in C_k$ e $R_j \in C_k$ if and only if for the same arrival and destination paths. With $S = \{1, ..., m\}$, $S \in \mathbb{R}^{m \times 2}$ the available London sensor set for determined pollution agent, so for each point $r_i \in R$ the pollution record for sensors $s \in S$, as shown in Algorithm 2.

The London Sensor Network API used for pollution factor attribution does not attend routes far from sensors tolerance in the built dataset. Similarly, other routes can contain readings from more than one sensor and consider an incoherent pollution level. For such cases, the proposed method excludes invalid routes for correct attribution and considers a tolerance `t = 300 m`.

Algorithm 2 assigns pollution levels to alternative route points for the given OD pair. The pollution record to each latitude-longitude pair attribute the pollution value, calculated from $pollution(s_j)$. After declaring C as the path set and S as the available sensor set, `line 1` defines an empty set for pollution associated with the determined path, filled with the pollution value from the nearest sensor. `lines 2` and `13` receive path length and starts sensor distance as 0. In `lines 4` to `6`, the algorithm starts the for-loop iterating $routePollution$ to every $routesPollution$ (path alternatives). In `lines 7` to `10`, a second for-loop inside the first one is initiated, iterating associated points near the path. `lines 11` to `15` begin the third loop iterating existing pollution sensor data for $pointPollution$. All loops finish attributing the values in `lines 16` to `21`. Finally, in `line 22`, the algorithm returns the pollution value for each route alternative in the given set ID. The algorithm returns 142 OD pairs with pollution-normalized values attributed, allowing the correct method application.

Algorithm 2 Routes pollution attribution and exclusion

Require: $C \neq \emptyset, S \neq \emptyset, t = 300$
1: $routesPollution = \emptyset$
2: $N \leftarrow length(C)$
3: $distance \leftarrow 0$
4: **for** $k \leftarrow 1$ to N **do**
5:     $routePollution = \emptyset$
6:     $minorDistanceRoute \leftarrow \infty$
7:     $M \leftarrow length(C_k)$
8:     **for** $i \leftarrow 1$ to M **do**
9:         $P \leftarrow length(S)$
10:         $pointPollution \leftarrow 0$
11:         $minorDistancePoint \leftarrow \infty$
12:         **for** $j \leftarrow 1$ to P **do**
13:             $distance \leftarrow haversine(r_j, s_j)$
14:             **if** $distance < minorDistancePoint$ **then**
15:                 $minorDistancePoint \leftarrow distance$
16:                 $pointPollution \leftarrow pollution(s_j)$
17:             **end if**
18:         **end for**
19:         $routePollution \leftarrow pointPollution$
20:         **if** $minorDistancePoint < minorDistanceRoute$ **then**
21:             $minorDistanceRoute \leftarrow minorDistancePoint$
22:         **end if**
23:     **end for**
24:     **if** $minorDistanceRoute > t$ **then**
25:         **return** $\emptyset$
26:     **end if**
27:     $routesPollution \leftarrow routePollution$
28: **end for**
29: **return** $routesPollution$

The Haversine function, shown in Equation 7 and `line 13`, calculates the distance between a tracepoint and a sensor considering the earth curvature. State-of-art solutions use geographic coordinates handling with this equation for the appropriate distance obtaining, represented by OD pair.

$$D = 2arcsin\left[\sqrt{sin^2(\frac{r_1 - s_1}{2}) + cos(r_1)cos(s_1)sin^2(\frac{r_2 - s_2}{2})}\right] \tag{7}$$

We compare the results from the four user profiles proposed to each greedy profile. The eight greedy profiles choose the alternative tuple with maximum value for only one criterion of all, as a common method in commercial navigation systems considers only Length, Traffic, and Duration when choosing a car trip. We measure the relation between every profile and the maximum value for a route choice with a mathematical method.

We use the Percent Deviation From a Known Standard (PDFKS) [Solé et al. 2022] for comparing the profiles to greedy profiles. The Equation 8 calculates the PDFKS value, where *p*[*f*] is the average value from feature *f* for each profile *p*, and *std*[*f*] is the best average value for feature *f* among all profiles results from the dataset values.

$$M_{pf} = \frac{p[f] - std[f]}{std[f]} \times 100\% \tag{8}$$

4.2. Results

Figure 3 shows the PDFKS matrix, where the *M* matrix rows represent the selection profiles (*p*), *i.e.,* four user profiles and eight greedy profiles, and the columns represent the trip features (*f*) for evaluation. The PDFKS metric value represents the best average value for each profile to the known standard (*std*[*f*]). For example, the first Safe profile cell shows a 3.6% increase to the best average value for crime, validating the Safe profile route selection with less crime than the other profiles after the onlyCrimes profile. Also, Safe presented a 39.5% deviation from the best average for accident feature, representing its second priority for route choice. This value does not represent the more significant performance for the accident due to the higher priority for crime feature and the data arrangement. In a feature less related to the main priority, the Safe profile scored a -46% deviation for the attraction feature, which its weight has less impact on selection.

In contrast, Nature, Attraction, and Traffic ratio have negative percentage PDFKS values in the matrix because the method searches for the higher nature and attraction occurrence and a higher traffic ratio that indicates a less congested road, implying on a raw value less than the known standard, resulting negative percentage. A lower raw value indicates the best route selection for all other features containing positive PDFKS values; the better selection method is with features closest to 0%.

We apply the absolute sum method for all 12 profiles, validating the user profiles, summing all elements without considering negative values. The lowest absolute sum of PDFKS for each profile represents the better selection method, considering all routes, as shown in Figure 4. We can note that greedy preferences have the known standard value (0%), indicating the best routes for a single feature, but tend to deviate more from all other features. Each proposed profile (Worker, Green, Safe, and Tourist) correlate to more than one feature, where we differentiate with colors the relationships and compare the resulting performance for all cases.

The Green profile has the closest value for the pollution standard and outranks the onlyPollution in other features and has the second best deviation from nature feature (-14.2%), resulting in a greener experience route. The Safe profiles outrank onlyCrimes (3.6%) and onlyTraffic (-1.7%), when the Crime feature is its higher priority, and higher traffic indicates slower paths and more dangerous routes, with the best deviation from crime. The worker profile overcomes its higher weighted features: onlyLength (6.4%), onlyDuration (1.9%), and onlyTraffic (-1.8%), surpassing the standard navigation systems in selection. The Tourist profile has better route selection than onlyAttraction, deviating

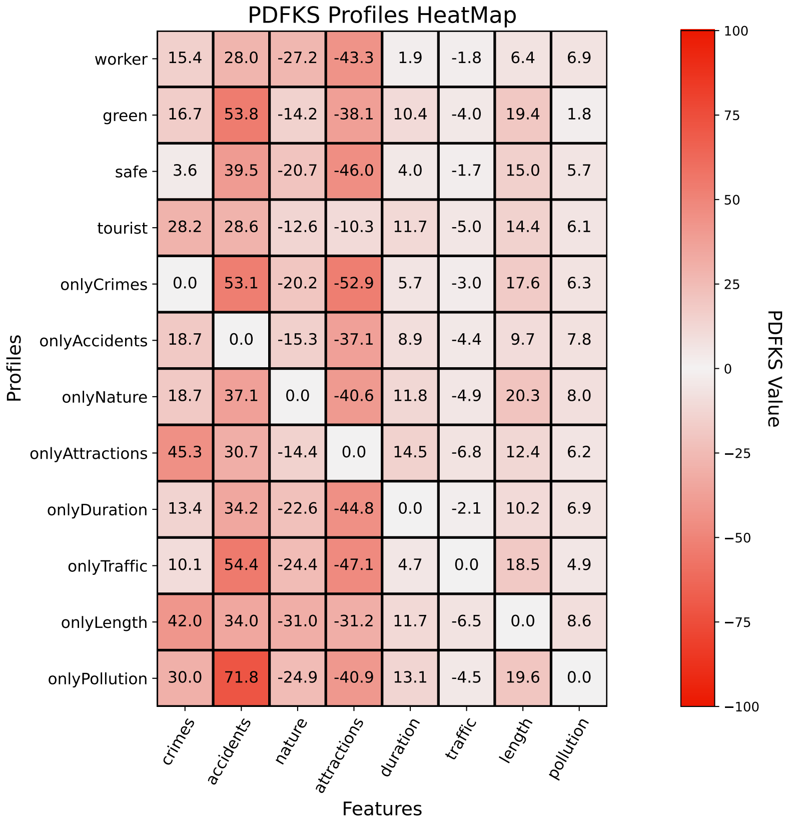

Figure 3. Methodology overview for route selection

from the standard attraction value with the best performance (-10.3%) and from nature feature (-12,6), with better performance for tourist users.

We note that all greedy profiles have the best performance for its features priorities but have a higher deviation in other features. The onlyAccidents have the lower absolute sum representing the best profile for selection in the evaluated environment. Otherwise, we note that the existence of an AHP profile with balanced weights for each criterion obtains better performance than single-criterion profiles, explaining the onlyAccidents higher performance. In other words, for diverse environments datasets, a specific profile with distributed weights, a priority for a few features, can outrank a greedy option.

In summary, our proposed user profile obtained a higher performance than its greedy opponents or was close to surpassing them. Tourist profile obtained the second (116,9%) best absolute sum and outranked the onlyAttractions greedy opponent by 13,4%, using the difference percentage. The Worker is the third (130,9%) best performance compared to greedy options, outranking them by 23,53% in the average difference percentage. Also, Safe and Green profiles surpassed their greedy opponents by 23,8% and 46,4%, respectively. This result indicates excellent usability for our method for considering all contextual data for selection than prioritizing only one criterion.

5. Conclusion

We propose a route selection method using multi-criteria decision-making and personalized routes for urban trip path choice considering all eight features. By comparing the greedy

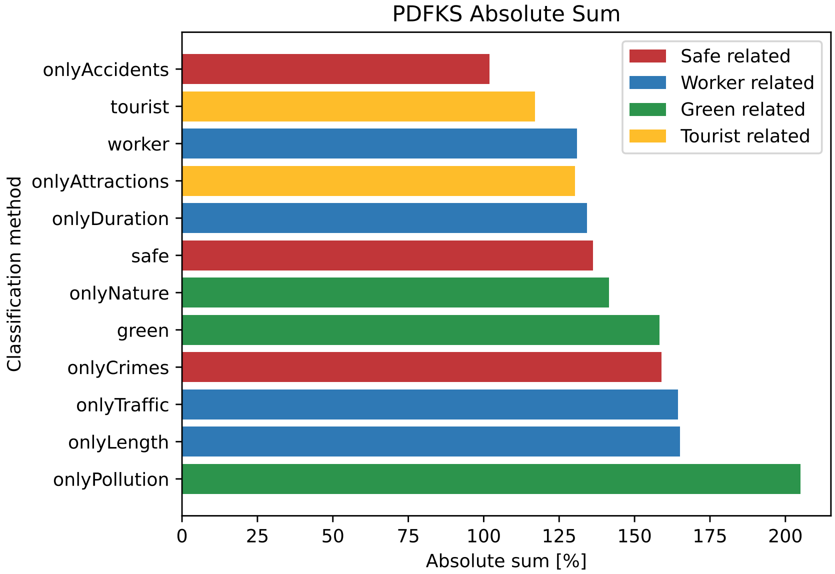

Figure 4. PDFKS absolute sum for each selection profile

profile for each feature, we observe a result closest to the best value from all features' best average value. This approach shows that traditional navigation systems can offer faster or healthier routes but can lead to dangerous or unpleasant paths. With the pollution factor addition, our method can prevent and alert the drivers and authorities to the polluted air threats, raising the quality of life. Furthermore, we developed the approach with simple mathematical methods for easy applicability in navigation systems.

In future works, a routing system can be built integrating various features and a multi-modal approach for intelligent public transportation. The system can consider IoT-enable feature prediction for real-time route selection, integrating user devices into a more extensive urban computing solution.

Acknowledgment

This work was conducted with partial financial support from the National Council for Scientific and Technological Development (CNPq), the Coordination for the Improvement of Higher Education Personnel (CAPES) - Finance Code 001, and Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) under grant 054/2021.

References

Bellini, P., Nesi, P., and Pantaleo, G. (2022). Iot-enabled smart cities: A review of concepts, frameworks and key technologies. *Applied Sciences*, 12(3).

Bibri, S. E. and Krogstie, J. (2020). The emerging data–driven smart city and its innovative applied solutions for sustainability: The cases of london and barcelona. *Energy Informatics*, 3:1–42.

Brito, M., Santos, C., Oliveira, H., Cerqueira, E., and Rosário, D. (2022). Air pollution calculation for location based social networks multimodal routing service. In *VI Workshop de Computação Urbana*, pages 280–293. SBC.

Han, X., Hu, X., Wu, H., Shen, B., and Wu, J. (2020). Risk prediction of theft crimes in urban communities: An integrated model of lstm and st-gcn. *IEEE Access*, 8:217222–217230.

HERE (2023). Build apps with here maps api and sdk platform access: Here developer.

Hsieh, H.-P. and Lin, F. (2022). Recommending taxi routes with an advance reservation – a multi-criteria route planner. *International Journal of Urban Sciences*, 26(1):162–183.

Kaivonen, S. and Ngai, E. C.-H. (2020). Real-time air pollution monitoring with sensors on city bus. *Digital Communications and Networks*, 6(1):23–30.

Kelly, F. J. and Kelly, J. (2009). London air quality: a real world experiment in progress. *Biomarkers*, 14(sup1):5–11. PMID: 19604051.

Legaspi, J., Bhada, S. V., Mathisen, P., and DeWinter, J. (2020). Smart city transportation: A multidisciplinary literature review. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 957–964.

Liu, D., Tang, J., and Zhang, C. (2019). Application of data analysis in trend prediction of different crime types in london. In Mao, R., Wang, H., Xie, X., and Lu, Z., editors, *Data Science*, pages 445–467, Singapore. Springer Singapore.

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., Niyato, D., Dobre, O., and Poor, H. V. (2022). 6g internet of things: A comprehensive survey. *IEEE Internet of Things Journal*, 9(1):359–383.

OpenStreetMap, c. (2017). Planet dump retrieved from https://planet.osm.org . `https://www.openstreetmap.org`.

Rodrigues, L. S., Sammarco, M., Detyniecki, M., and Campista, M. E. M. (2021). CRAWDAD dataset ufrj/london-trajectories (v. 2021-03-22). Downloaded from `https://crawdad.org/ufrj/london-trajectories/2021-03-22`.

Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1):9–26. Desicion making by the analytic hierarchy process: Theory and applications.

Sarraf, R. and McGuire, M. P. (2020). Integration and comparison of multi-criteria decision making methods in safe route planner. *Expert Systems with Applications*, 154:113399.

Savithramma, R. M., Ashwini, B. P., and Sumathi, R. (2022). Smart mobility implementation in smart cities: A comprehensive review on state-of-art technologies. In *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pages 10–17.

Solé, L., Sammarco, M., Detyniecki, M., and M. Campista, M. E. (2022). Towards drivers' safety with multi-criteria car navigation systems. *Future Gener. Comput. Syst.*, 135(C):1–9.

Wu, M.-Y., Ke, C.-K., and Lai, S.-C. (2022). Optimizing the routing of urban logistics by context-based social network and multi-criteria decision analysis. *Symmetry*, 14(9).

Zhang, D.-G., Wang, J.-X., Zhang, J., Zhang, T., Yang, C., and Jiang, K.-W. (2022). A new method of fuzzy multicriteria routing in vehicle ad hoc network. *IEEE Transactions on Computational Social Systems*, pages 1–13.