International Journal of Science Academic Research
Vol. 06, Issue 02, pp.9352-9362, February, 2025
Available online at http://www.scienceijsar.com

ISSN: 2582-6425

Research Article

A SYSTEMATIC REVIEW OF THE ESMVERE ALGORITHMS, WHICH ARE CONTRIBUTIONS IN THE FIELD OF UNCAPACITATED FACILITY LOCATION PROBLEMS AND K-MEDIAN PROBLEMS

***Emmanuel Siyawo Mvere**

Mechanical and Production Engineering Department, University of Mauritius, Moka, Mauritius

Received 25th December 2024; **Accepted** 20th January 2025; **Published online** 28th February 2025

Abstract

The objective of the present paper is to review my personal contributions in the field of Uncapacitated Facility Location Problems. This paper is an analysis of scientific contributions, mainly the relevance, originality and significance of three published articles, around location science. The published journal articles are derived from a single source asnovel solutions to a real-world case study. They contrast and revolutionize over 60 years of solutions to the UFL problem and k-median problem. The research introduces a new data analysis method, a Hamming distance-based k-median clustering algorithm based on the R Language. The research also introduces novel reformulations of the UFL and k-median problems which replace traditional distance matrices with Global Position System (GPS) based Euclidean Distance calculations, simplifying data input and reducing computational complexity. The reformulations are based on the CPLEX Opl language. The novel software inventions are all used to solve a real-world problem which involves the near optimal siting of Hazardous Waste Used Lead Acid Battery (ULAB) collection facilities in the republic of Mauritius. The thrust of this review is to analyze the novelty, compare, contrast and discuss all the remarks made by various reviewers. Hence this review combines and summarizes the three papers and explains the contribution of the three novel software inventions developed to solve a real-world problem.

Keywords: UFL problem, k-median problem, p-median problem, TSP problem, Hazardous Waste

1. INTRODUCTION

[1][2][3] explore reverse supply chain network design problems in the automotive industry of the republic of Mauritius, specifically focusing on the challenge of recycling Used Lead Acid Batteries (ULABs). The authors examine logistical issues related to the Uncapacitated Facility Location problem (UFL) and the k-median location problems. Thus [1][2][3] deals with the important problem of improving the safety and effectiveness of the collection, processing, and disposal of one of the most hazardous parts of the vehicles, the ULAB.

1.1 An Analysis of the Used Lead Acid Battery (ULAB) Collection Network Mauritius Using the ESMVERE- R-Hamming-k-median Clustering Method [1]

The primary contribution of [1] is its investigation of the important and complex logistical problem of ULABs, which are a hazardous waste problem [4], faced not only by Mauritius, but also by other countries. Used batteries, in general, particularly the lead acid ones are among the most hazardous waste, and an effective collection, processing, and disposal of them is of paramount importance to the health and safety of all communities, particularly those that are more developed, for example communities that have more vehicles per capita [3]. The life of these batteries is shorter than the life of vehicles that may undergo a few battery replacements in their usable life. This is a highly relevant problem, given the potential environmental impact of discarded batteries. [1] deals with quantitative research where a survey was conducted among a sample of industry experts.

Corresponding Author:** ***Emmanuel Siyawo Mvere,
Mechanical and Production Engineering Department, University of Mauritius, Moka, Mauritius.

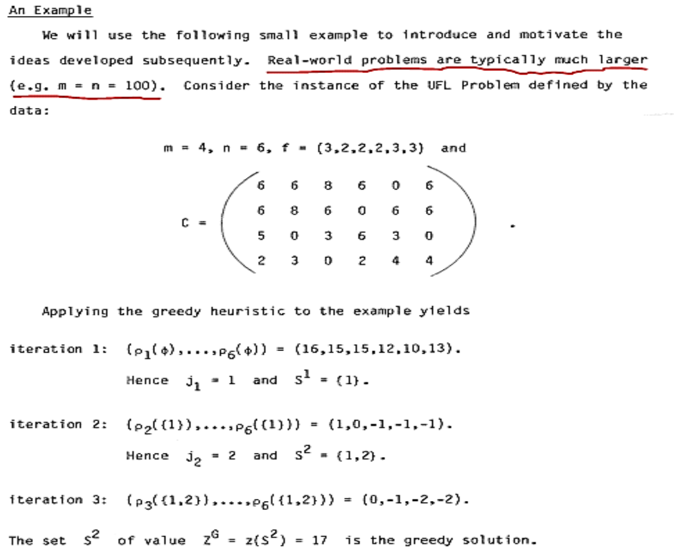
A total population of 137 experts who were directly and indirectly involved in the trade of ULABs, that is importation, reselling, repair and collection were selected. This included relevant members of the environmental and human health ministries. A sample population of 40 experts was selected [1] to represent the total population of 137. A questionnaire was drafted which contained about 20 questions (some open ended) which sought to describe the existing structure and network of the ULAB life in the republic of Mauritius, to determine the impact of the current trade and disposal system. A descriptive statistical analysis of the normative answers and briefs of responses to open ended questions were given, which showed that many of the respondents acknowledged the need for a new and safer collection system for ULABs. The small size of the population and sample in this specialized area is justified given that Mauritius is a small island which hosts a total population of less than one million two hundred thousand people, 1.2 million people. [1] Generated some commendable insights on the yearly system of disposal, effectively correlating it with environmental degradation and societal health risks. Thus [1] acknowledges the wide gap in research on this area in the country Mauritius, where the last research on lead poisoning was conducted in 1995 by [5]. Thus, despite the adverse dangers on lead poisoning, no research on the topic had been conducted in Mauritius for almost 30 years. The second contribution of [1] is the introduction of a novel data analysis method which was used to analyze the data. While the research describes the analysis of a questionnaire dataset, it contributes a new software invention, a novel clustering method which was used to identify and interpret the clusters of the respondents. The new method ascribes a numerical value to the best choice of the number of clusters k, when using the k-median hamming distance variant [1]. Thus, this is usually the most complex decision when clustering data using either the k-means or its variant the k-median [1]. There is a need to be sure of the optimum number of clusters that give the best value of the k for the best clustering performance [1]. Thus [1] presents a new data analysis method that may be used to analyze questionnaire responses in many other contexts.

1.2 [2][3] The Near Optimal Siting of Hazardous Waste (Used Lead Acid Battery (ULAB)) Collection Facilities in the Republic of Mauritius Using the [2] ESMVERE CPLEX UFL Problem Solver , [3] ESMVERE CPLEX k-median Problem solver algorithm

[2][3] presents a mathematical framework that consists of two facility location problems the UFL, and the k-median problems as alternative approaches to finding the best location sites for the collection facilities of the ULAB batteries. [2][3] present an innovative approach to the siting of hazardous waste collection facilities in Mauritius using the novel ESMVERE Cplex UFL problem solver algorithm, and the ESMVERE Cplex k-median problem solver algorithms. Both papers address a significant gap in the literature by applying UFL and k-median problem-solving techniques in a real-world setting, using the actual data from the Mauritius statistical office. This addresses a wide gap in research for these problems as all known prior work on these problems was mostly theoretical improvements of the UFL and k-median approximation algorithms [6][7][8][11].

[6][7][8] solve the problem theoretically by developing faster algorithms which are seldom tested on software generated data. Thus, they would achieve theoretically faster algorithm solutions, however in most cases these would be unusable [7] in the real world. This would explain why it is still rare to find a "solved" real-world case study in this area of research. Although [8] propose an algorithm that is aimed at attempting real world cases, there is no illustrated example of this achieving the proposed results. All the solutions and improvements would not provide a solution that the reader could test and use in their own context thus [2]and [3] closed this gap. Both [2]and [3] propose a realistic usable solution that "solved" this problem in Mauritius to site ULAB collection facilities. This work demonstrated fluently how the UFL problem can be solved on real data. Both [2] and [3] illustrated a step-by-step instance creation process which familiarizes the reader with how to solve a realistic real-world problem. In the Appendix both [2] and [3] illustrated an unrelated, independent, separate sample solution with results, and provided the code for the readers to test and develop in their own context. [2] [3] clearly demonstrated the algorithm solution to the real-world data from Mauritius using the ArcGIS mapping. The solutions are also visible when they are input on google earth.

The primary contribution of both [2] and [3] is the removal of the distance matrix as input in the calculation for the UFL problem and the k-median problem, replacing the distance matrix with the GPS input, Euclidean distance formular. This simplified the data input and reduced computational complexity for the algorithm making it easy, usable and convenient as a solution. The small illustration in the appendix [2][3] proves that the solution is usable in many other cases, meaning that the same algorithm can be used for other similar problems. Since their introduction over 60 years ago [12][25] both the UFL and k-median problem have been solved using the distance matrix. [9][10] are examples of the CPLEX Opl software algorithms that solve the UFL problem showing the input of the distance matrix. [13] is a p-median example of how both problems were solved. Thus [2] [3] introduce a solution that does not need the input of a distance matrix but uses GPS based Euclidean distance calculations. [9][10] [11][13] and all known prior solutions used the input of a distance matrix in the calculation of the UFL problem and the k-median problem.


An Example

We will use the following small example to introduce and motivate the ideas developed subsequently. Real-world problems are typically much larger (e.g. m = n = 100). Consider the instance of the UFL Problem defined by the data:

$m = 4,\ n = 6,\ f = (3,2,2,2,3,3)$ and

$$C = \begin{pmatrix} 6 & 6 & 8 & 6 & 0 & 6 \\ 6 & 8 & 6 & 0 & 6 & 6 \\ 5 & 0 & 3 & 6 & 3 & 0 \\ 2 & 3 & 0 & 2 & 4 & 4 \end{pmatrix} .$$

Applying the greedy heuristic to the example yields

iteration 1: $(\rho_1(\phi),\dots,\rho_6(\phi)) = (16,15,15,12,10,13)$. Hence $j_1 = 1$ and $S^1 = \{1\}$.

iteration 2: $(\rho_2(\{1\}),\dots,\rho_6(\{1\})) = (1,0,-1,-1,-1)$. Hence $j_2 = 2$ and $S^2 = \{1,2\}$.

iteration 3: $(\rho_3(\{1,2\}),\dots,\rho_6(\{1,2\})) = (0,-1,-2,-2)$.

The set S^2 of value $Z^G = z(S^2) = 17$ is the greedy solution.

Figure 1. UFL Matrix solution [11]

It would be highly difficult to use this solution in a real-world scenario, which could explain why there are not so many real-world examples. All known prior innovations to the UFL and k-median problem solutions were centered on improvements and innovations around the matrix [11] [13]. Thus [2][3] propose a solution which is a novel breakthrough spanning over 60 years [11] of work in this area.

1.3. Contribution Of The Euclidean Distance Formular To The UFL and k-median Problem In CPLEX Opl [2] [3]

To the best of our knowledge, all known prior algorithms use the input of a distance matrix, as input in the calculation. The removal of the distance-matrix in the calculation is the main difference between the model solution developed to solve the UFL and k-median problem in [2][3] compared to all known prior methods solving this problem. For example, [9] uses the distance matrix below (a), in a small nine warehouse example. Even in their short example, to ensure that the distance of each point to the other remaining eight points is combinatorically considered, is a big combinatorial task [2][3]. This becomes even more complex when you must gather the data between all possible combinations of distances, for example if there are 50 locations : From the first location we label it L 1 to the last location L 50

L 50 ->2^{49} lists all the possible pairs
L 49 -> 2^{48} lists all possible pairs (same as above, excluding L 50)
$2^{49}+2^{48}+\ldots+2^{0}=2^{50}-1$

Continue with this to the last location L1 then add them up we have 1.1258999e+15 location distances to place in a distance matrix. This is somewhat easy when it is randomly generated by software, however, in real world practice this is extremely complex. Below is a distance matrix of only 9 locations [2][3].

Distance Matrix [9][2][3]

$$Distances\ [\,[0,720,790,297,283,296,461,796,996]\\ [720,0,884,555,722,461,685,245,1099]\\ [790,884,0,976,614,667,371,645,219]\\ [297,555,976,0,531,359,602,715,1217]\\ [283,722,614,531,0,263,286,629,721]\\ [296,461,667,359,263,0,288,479,907]\\ [461,685,371,602,286,288,0,448,589]\\ [769,245,645,715,629,479,448,0,867]\\ [996,1099,219,1217,721,907,589,867,0]]; \qquad (a)$$

This would be impractical in many cases in the real world where there might be hundreds or thousands of location points, and the distances that make each pair of distances [2][3]. Thus, a new model is developed that does not need the input of the distance matrix but uses the formular for the Euclidean distance between two points and the input of actual GPS coordinate location points to solve the k-median problem [2][3] .

The formular for the Euclidean distance between two points is:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (b)$$

The UFL,k-median Cplex Opl Euclidean distance formular developed for this study

$$function\ getDistance\ (ulabcity1, ulabfacilitycity1)\ \{\\ return\ opl.sqrt(opl.pow(ulabcity1.x\\ - ulabfacility1.x, 2)\\ + opl.pow(ulabcity1.y\\ - ulabfacility1.y, 2))\ \} \qquad (c)$$

The new method solves the UFL and the k-median problem by making a choice between facility locations and client locations, without considering the distance between all the points as a distance matrix. The newly developed method only requires the GPS coordinate location points [2][3]. This means that it opens the best out of *n* potential facility locations and assigns only *m* clients to these open facilities per calculation [2][3]. They are both heuristic solution techniques, Integer binary linear programs with binary variables that solve the UFL and the k-median problem. The new method could also be a useful contribution to UFL and k-median approximation algorithms, where they could contribute to faster approximation algorithms for the UFL and the k-median problem.

2. REVIEWER DISCUSSIONS

2.1a Reviewer 1 comments[2][3]

"The originality of the work is limited, as the models (k-median and Uncapacitated facility problems) and methods (direct solution of the models using the OPL-Cplex tool) addressed are widely known in the literature of Operations Research. Indeed, it is not difficult to find material about these problems (see, for example, https://perso.ensta-paris.fr/~diam/ ro/on1ine/cptex/cplexl271/OPLStudio/opl quickstart/ topics/ opl_quickstart_pmedian.html for the k-median and https:// github.com/samarthrnistry/Optimization-Facility-location for the Uncapacitated facility 1ocation). The contribution of the work is mainly in addressing a specific problem in Mauritius with these known tools."

Discussion

Research on these problems is on-going. The reviewer's comments present some confronting arguments that demand an analysis. A brainstorm of what the author immediately construed is:

- The reviewer might mean the scientific community must not improve solutions to these problems any longer. The international scientific location science expert community must halt, refrain and stop all its current research on these problems?
- The reviewer might mean all the methods used to solve these problems are the same. Is it possible that the custom new solutions proposed can fit under a blanket term of well-known tools?
- The invented solutions are in the same class as Mathematical model problems or the software language used to develop them, and not on how they independently and uniquely solve these problems.
- Originality of research is based on the model problems UFL, k-median and not on solutions developed to solve the problems.

Indeed, both the UFL and the k-median problems are widely known in literature, since they have been in existence since [12][25]. However, as many solutions may exist, there is always need for a better solution. A small improvement can make a significant difference. More so, the hardness of approximation on both these problems has not been achieved 1.463 for the UFL problem [21] and 1+2/e for the k-median problem [22]. Thus, the quest is to develop an original, novel and better solution than what is already well-known. This surely cannot be overlooked.

"It can be seen that researchers have tried different permutations and combinations of various exact, heuristic and metaheuristic algorithms to develop better solution methods than the existing ones" [24].

Open minded academia's will ask relevant and crucial questions such as: How does the new solution change the paradigm? How does it challenge the existing solutions? where does the new and innovative solution lead to? how does it improve our society?

The statement "originality of the work is limited" connotes to the work having existed before. It is true to say the UFL and k-median problem has been in existence [12][25]. This statement, however, in this context, points to imply that the work has proposed new problems, which are the UFL problem and the k-median problem. This is not the case. The work has clearly proposed new solutions to these existing problems. These solutions are based on a breakthrough development in how the problems are solved. Which could in turn lead to the new solution becoming the best solution to use in this case of study, and in many similar cases, based on the results and evidence of its success.

"methods (direct solution of the models using the OPL-Cplex tool)" The method used is an integer, binary heuristic solution algorithm which was developed firsthand, it is a newly invented algorithm, however, the reviewer implies from the statement that it is a pre-configured OPL-Cplex tool which was in existence. This statement appears like an accusation, it is fraudulent. Then comes the evidence supplied [10] in this case Figure 3, which is in no way related to how the solution [2][3] solves the problem.

Type	Definition	Code availability
Proprietary software	Software whose use or redistribution is restricted in the way that users cannot do it without permission	Not publicly available
Custom software	Software developed for one organisation or company	Often not publicly available
Open-source software	Codes that anyone can inspect, apply, modify and enhance	Publicly available

Figure 2. Custom software [33]

“Some researchers have written custom software that directly structures and solves location models [34][35], relying on general optimization solvers. While custom software allows developers to specify their models, it comes with obvious shortcomings. It is challenging to develop new software, which requires specific technical skills software design, tackling data input, model implementation and integration with solvers. Moreover, in academia, researchers generally neither disclose the technical details nor release the code of developed software”[33]

The problems are not original, but a better solution which exceeds the world’s best in this area might be a new, original and better method. In an academic setting, the reviewer is also given a Turnitin score. The Turnitin score defines how much of the work has been derived from uncited sources, or whether the work has been plagiarized from another source. If the paper is submitted with a high Turnitin score, then that is enough evidence that the work is not original. However, if the Turnitin score is fairly below an acceptable number, then that is evidence that the work may be original, such as in this case, the Turnitin score is below 10.

Otherwise, there is a need for evidence or full disclosure of similar work. Evidence of similar work. A full year may be given to the reviewer to go through the body of knowledge in this field, this is enough time for the examiner or reviewer to find if similar work has existed. It is damaging, unprofessional and incompetent to claim a newly developed custom solution is a well-known tool. This statement oppresses, disqualifies the newness of a custom solution. It is highly unprofessional to mention anything at all without detailed evidence. Figure 3, [10] is not evidence related to Appendix A on both [2][3] in any way. In a high school examination such errors do not exist. A high school student would never agree that [2][3] and [10] are similar programs. See Appendix A and B

2.1b Reviewer 1 comments

The originality of the work is limited, as the models (k-median and Uncapacitated facility problems) and methods (direct solution of the models using the OPL-Cplex tool) addressed are widely known in the literature of Operations Research.

Discussion

Research work on the UFL problem and the k-median problem is ongoing.[8][14][15][16][17][18][19][20] are all recently published improvements on both problems in this area of research. The demand for work on these problems proves even more that they are one of the most relevant areas of research in this day. The approximability lower bound of both problems which is the level of hardness of the problems, have not yet been achieved 1.463 for the UFL [21], for the k-median 1+2/e [22]. Recent increases in work on these problems prove even more the usefulness of solutions to these problems. With the advent of Artificial Intelligence (AI) and Machine Learning (ML) more advanced modern-day uses of the algorithms that solve these problems have arisen. Self-driving vehicles, drones, computer networks, pattern recognition, image processing, social media, computer vision, data mining, all these areas of application and more, have developed the necessity to know more about the efficient allocation of resources. An example is, we might need to know where the median of the locations is to process efficient deliveries for self-driving vehicles, drones etc. [2][3] go many steps further by allowing the use of GPS coordinates for ArcGIS or Google earth in their calculation. The [2][3] solutions are novel and unique; they solve these problems in a unique way.

The problems have existed since [12][25] but each solution in [8][14][15][16][17][18][19][20] addresses the problems differently. They are published by the best journals in the world, after a highly credible review process, because they are uniquely developed solutions that extend the body of knowledge in the area of research which is ongoing. This is one of the most relevant areas of research today in terms of the demand for solutions to the k-median and the UFL. Many cases demand the elimination of outliers. [8][14][15][16][17][18][19][20] are multiple independent software solutions to the UFL and k-median problem, these solution algorithms are developed in multiple independent programming languages. IBM ILOG CPLEX Optimization Programming Language is one such Language that can be used to write a solution to the UFL or k-median problem. Other Languages include Gurobi, Python, MATLAB. The choice of which optimization programming language is used depends on the choice and familiarity of the developer with a language. [8][14][15][16][17][18][19][20] all solutions developed in various optimization programming languages however they are not similar in anyway. [6][7] are among the best works in this area of research, they can be termed legendary, both authors clearly mention they were developed using the CPLEX OPL. To a human of average intelligence, the two solutions are so contrasting and nowhere near similar that they fit under a blanket term of well-known CPLEX tool. It would be absurd for the award of the best student paper at FOCS and ICALP [6].

The same goes for [2][3] these solutions are independently and uniquely developed because they are the first known solutions of these problems in which the distance matrix is removed. To prove this difference the example given by the reviewer uses a distance matrix Figure 3 [10]. While it is clearly not difficult to find material about the problem and solutions developed in CPLEX OPL, each solution developed is unique. A careful and wise analysis of every solution and how it permeates the boundaries of existing knowledge is always necessary. Comparing [2][3] and the examples given by the reviewer [10] in this case would clearly show the difference between these two solutions. Whereas [2][3] solves the problem with the input of GPS coordinates and Euclidean distance-based calculations, [10] uses the input of a distance matrix as is the norm for the past 60 years. Hence [2][3] is a significant step because it solves the problem efficiently affording the results. It is developed in such a way that makes it easier for the reader to understand. However, a simple analysis by an entry level high school computer programming student would point out the difference of the programs between [2][3] against [10] fluently. An entry level high school programming student understands that a program consists of input, processing and output. The first question for an entry level high school student (and even some brilliant primary school students) would ask is, where is the input? Or what is the input? Just those two questions would lead to the fascinating revelation that there is a big difference between [2][3] and [10].

2.1c Reviewer 1 comments

Indeed, it is not difficult to find material about these problems (see, https://github.com/samarthrnistry/Optimization-Facility-location for the Uncapacitated facility 1ocation). The claims about consistent contributions to the OR 1iterature should be toned down as the work does not introduce any new methodologies or techniques to the scientific community.

Discussion

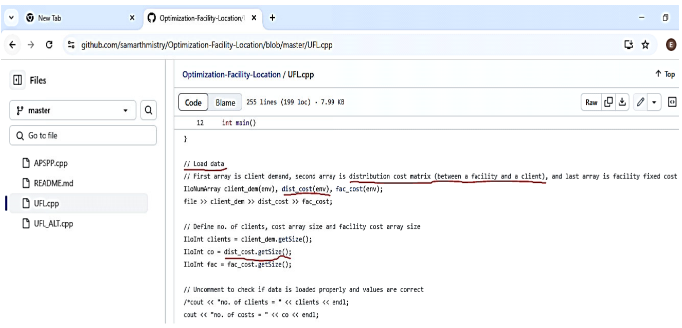


Figure 3. CPLEX OPL tool [10]

The underlined code commenting sections clearly show that the program is going to load data, and it is written that the "second array is distribution cost matrix (between a facility and a client)". Thus, the main contribution of [2] and[3] is the elimination of this distance matrix. To the best of our knowledge all known prior solutions used the input of a distance matrix [9] [10][11] [13]. The main contribution had been stated as a novel algorithm that solves the UFL problem, which in all simplicity, a knowledgeable reviewer would have analyzed this difference and argued around, had it been done anywhere in the world before the review? Thus, leading to an investigation which involves 1) checking the Turnitin score, and 2) conduct a google search. In the modern era of AI, one would have come up with an algorithm closer to the original [2] [3] if it had existed. A review of these simple 2 steps would have proved a person of wise and dignified intellect who does not rush in making claims without a detailed analysis.

2.1d Reviewer 1 comments

(see, for example,https://perso.ensta-paris.fr/~diam/ro/ on1ine/cplex/cplexl271/OPL_Studio/opl_quickstart/ topics/ opl_quickstart_ pmedian.htm1 for the k-median)

Discussion

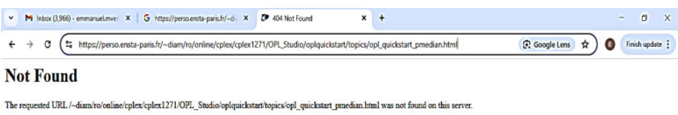


Figure 4. NonexistentCPLEX OPL tool [10]

The website referred for does not exist??

2.1e Reviewer 1 comments

The contribution of the work is mainly in addressing a specific problem in Mauritius with these "known tools".

Discussion

The sections above clearly spelled out the difference between [1][2][3] and what is "known" in this area of research. An expression of this in the simplest of terms is, the UFL and the k-median problems are like the question paper in an examination, and the solutions to these problems solutions such as [1][2][3] [6][7][8][9][10][11][12][13] [14][15][16][17] [18][19][20] are the answer sheets. Hence the evaluation in all the best journals in the world and all the best Universities, is based on how the answer sheet extends what is already "known" or what is "widely available". How do the proposed solutions appear newer or better than all "known" prior solutions to these problems. The evaluation is not based on the (question paper,) which is existing knowledge but on how the (answer sheet) compares to what is already known. Hence if a solution is evaluated and is found to have existed or there being similar work then it is a "known tool", there should be evidence of a similar solution, in terms of input, computation time, speed, mathematical model, in a few cases the results. If none of these things are similar except the problem being solved and the language used, then these are two separate independent solutions such as [2][3] against [10].

1. There is the "model" (UFL problem and k-median problem [43])
2. Then there is the mathematical concept behind the solution the "maxim", (integer binary heuristic algorithm [2][3], simple multi-wave algorithm [20], primal-dual approximation algorithms [6], a hybrid tabu search with path-relinking simheuristic approach[8])
3. Then there is the software programming language used to develop an algorithm, which is the practical application of the mathematical concept.

A solution can be called similar if all 1,2,and 3 are the same. The solution [10] is not an integer binary heuristic solution, nor is it a simple multi-wave algorithm, nor is it a primal-dual solution. It appears to be a Mixed Integer linear program(MIP) or an Integer Programming Formulation. The reviewer is analyzing 1 and 3 for the review. 1 , "which is taught in basic causes, and is available everywhere"[43]. The Reviewer failed to distinguish between the three, 1, 2 and 3.

2.1.1a Reviewer 1 comments

However, the modelling and techniques adopted in the work are quite basic, as they are based on the direct solution of well-known mathematical models using the OPL tool.

Discussion

The removal of the distance-matrix in the calculation is the main difference between the model solution developed to solve the UFL and k-median problem in [2][3], and all known prior methods of solving this problem. It cannot be termed "basic" to develop invent a solution to a problem that has existed for over 60 years. A careful analysis of [2][3] will reveal how the solution was invented, this was not a preconfigured tool that was in existence, because every solution before [2][3] uses a distance matrix in the calculation. [2][3] is an everyday usable solution which allows the input of GPS coordinates in the calculation of this problem. Thus, again the statement "quite basic" should follow with evidence of a solution which is more advanced solving the problem. The approach in the examples given [10] use a distance matrix. Figure 3 is not solving this problem faster or more conveniently than [2][3]. The mathematical models are well known however the best student papers are being awarded on innovations and improvements to these models [7][8][14][17][19].

2.2a Reviewer 2 comments [2]

The paper examines the challenge of hazardous waste management from Used Lead Acid Batteries (ULABs) in Mauritius, a pressing public health and environmental concern due to the risks of lead exposure. It proposes a heuristic solution based on the Uncapacitated Facility Location (UFL) problem; a mathematical framework designed to minimize facility setup and transportation costs while ensuring adequate service coverage. The study introduces the ESMVERE Cplex UFL Solver, which replaces traditional distance matrices with GPS-based Euclidean distance calculations, simplifying data input and reducing computational complexity. Drawing on methodologies from related optimization problems, including the Traveling Salesman Problem (TSP) and p-median models, the authors develop an algorithm tailored for small- to medium-scale applications. Applied to real-world data from Mauritius, the model identifies 19 optimal ULAB collection sites, achieving cost-effective logistics and minimal transportation distances. The results, visualized with ArcGIS, align closely with geographic realities, underscoring the algorithm's practical applicability. However, the solution's scalability is limited, as it can process only up to seven facilities and 125 clients per iteration. Enhancements to address larger datasets are recommended for future research.

Discussion

The initial software used to develop the algorithm developed for this study had a limitation on instance input size, it was only reliably usable in cases where there were 7 Facilities and 125 clients [37]. This meant that it opened the best out of 7 potential facility locations and assigned only 125 clients to these open facilities per calculation. This limitation was with IBM ILOG CPLEX Optimization Studio 22.1.2 community edition. This is a free version of the software. The commercial version of the software does not have the same size limitations. The IBM ILOG CPLEX Optimization Studio 22.1.2 Developer Edition was used to test and compare the performance of the algorithms in larger instances. This enabled the algorithm to solve larger instances of sizes, up 1000 facilities and 3000 clients.

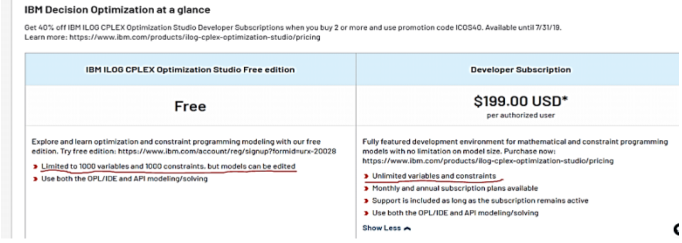


Figure 5. CPLEX Opl Size Limitation [37]

2.2b Reviewer 2 comments

This study's key contribution is the practical adaptation of theoretical UFL frameworks to address environmental challenges, offering a replicable model for hazardous waste management in other contexts. It highlights the potential for mathematical modeling to drive practical solutions to global environmental challenges as well. The paper addresses the pressing environmental and public health challenge of managing hazardous waste from Used Lead Acid Batteries (ULABs) in Mauritius. It proposes a heuristic solution grounded in the Uncapacitated Facility Location (UFL) problem, a mathematical model designed to minimize facility setup and transportation costs while ensuring effective service coverage. The use of GIS mapping further enhances the study's relevance, making the findings accessible to stakeholders such as policymakers and environmental agencies. While the paper provides a replicable framework for hazardous waste management, its scalability is limited to seven facilities and 125 clients per iteration, restricting its application to larger or more complex scenarios. Enhancing computational capacity and addressing larger datasets in future research could significantly extend its impact.

Discussion

This limitation was with IBM ILOG CPLEX Optimization Studio 22.1.2 community edition, Figure 5 [37] . This is a free version of the software. The commercial version of the software does not have the same size limitations. The IBM ILOG CPLEX Optimization Studio 22.1.2 Developer Edition was used to test and compare the performance of the algorithms in larger instances. This enabled the algorithm to solve larger instances of sizes, up 1000 facilities and 3000 clients. This challenge was addressed in [2][3]

2.2c Reviewer 2 comments

The paper addresses a critical environmental and public health issue in Mauritius by tackling the challenges associated with Used Lead Acid Batteries (ULABs). It offers actionable insights and recommendations for hazardous waste management, supported by real-world data and GIS mapping, with potential applications to other regions or waste types. However, the study is tailored to Mauritius only, and while the findings are valuable for the local context, there is limited discussion on adapting the model for diverse geographic or logistical conditions. A broader framework addressing these aspects would enhance its utility for global applications.

Discussion

Appendix A [2][3] hosts the algorithm, this model is independent and separate from the data. A thorough analysis of the code in Appendix A will reveal that the algorithm can be used in all global contexts, this will mean renaming some of the variable names to suit the chosen area of application. This in my view makes it usable globally. Appendix B on papers [2][3] hosts a small more detailed demonstration to which a google earth map is used to visualize the median location. This is illustrated so that it is easier to adopt the same algorithm in global application. The problem in this case applied specifically to Mauritius a small island where no such research on ULABs had been conducted in over 30 years. However, the solution developed may be applied to other countries as well.

2.2d Reviewer 2 comments

The paper does not sufficiently compare the proposed algorithm with state-of-the-art methods in terms of computational efficiency and solution quality. Including such comparisons would provide a clearer perspective on the model's relative strengths and weaknesses.

Discussions

The MED set of instances from [8] [23] [15] [20] were compared and used to analyze the quality and performance of both the ESMvere UFL and k-median algorithms on the larger cases[2][3] The sets contain 18 large-scale problems used by [15] in the context of p-median[2][3]. The tests proved that algorithms on both [2][3] are evidently and exceptionally faster than all the best in literature at solving these test instances.

Table 2

Instance	Gurobi	ILS		MDAS		HYB		ESMVERE CPLEX k-median		
	OPT	AVG	CPU	AVG	CPU	AVG	CPU	AVG	CPU (*optt*)	CPU(*time*)
500-10	798577	799077.1	46.89	798577.0	28.0	798577	33.2	798277.0	25.488	6.22
500-100	326790	327007.6	18.54	326922.375	25.6	326805.4	32.9	226798.4	11.402	19.49
500-1000	99169	99172.6	14.57	99196.0	22.4	99169	23.6	89170.9	17.13	47.16
1000-10	1434154	1435918.6	54.94	1434171.0	130.7	1434185.4	173.9	1334154	12.41	38.75
1000-100	607878	608202.3	160.76	607992.563	106.5	607880.4	148.8	597883.9	63.21	134.62
1000-1000	220560	221260.9	175.71	220626.563	84.4	220560.9	141.7	210585.2	139.42	273.47
1500-10	2000801	2002874.5	178.66	2000854.14	331.1	2001121.7	347.8	1800828.8	60.89	111.23
1500-100	866454	867654.7	385.46	867149.69	293.6	866493.2	378.7	746490.1	176.11	314.24
1500-1000	334962	338046.1	381.83	335400.813	218.7	334973.2	387.2	125059.1	251.28	702.70
2000-10	2558118	2559611.3	224.77	2558121.5	687.4	2558120.8	717.5	1458119.4	84.39	277.50
2000-100	1122748	1125471.9	980.89	1123936.5	562.3	1122861.9	650.8	1022854.5	241.26	660.59
2000-1000	437686	443025.7	997.64	438263.0	425.9	437690.7	760	337813.6	457.60	1201.25
2500-10	3099907	3107032.5	710.01	3100174.5	1116.7	3100224.7	1419.5	2100155.9	149.50	531.12
2500-100	1347516	1350446.6	1903.79	1348713.25	870.5	1347577.6	1128.2	1247566.8	711.44	1036.96
2500-1000	534405	540365.2	1940.58	535134.938	675.3	534426.6	1309.4	434593.6	1423.54	2737.16
3000-10	3570766	3579295.7	2605.87	3570820.75	1667.0	3570818.8	1621.1	2570766	187.43	876.77
3000-100	1602154	1607502.6	2987.58	1605083.63	1349.5	1602530.9	1977.6	1402512.2	889.13	1711.03
3000-1000	643463	652092.7	2996.92	644376.25	1017.3	643541.8	2081.4	543885.3	1522.41	3389.68

Figure 6. Computation efficiency[2][3]

2.2e Reviewer 2 comments

The model is designed for static scenarios and does not account for dynamic changes, such as fluctuations in waste generation or client locations over time. Incorporating a dynamic component could improve its adaptability and relevance in changing environments. By addressing these limitations, the paper could significantly enhance its impact and applicability, making it a more robust tool for hazardous waste management.

Discussion

Future work in this area, on these problems will include a study on Mobile Facility location Problems [43] and Dynamic Facility Location Problems

2.3a Reviewer 3 comments [2]

This paper presents an innovative approach to the siting of hazardous waste collection facilities in Mauritius using a novel ESMVERE Cplex UFL problem solver algorithm. It addresses a significant gap in literature by applying UFL problem-solving techniques in real-world settings. The study is well-structured and provides meaningful insights into reverse logistics for hazardous waste.

Strengths and weaknesses

Strengths:

1. Novel approach integrating UFL problem with real-world data.
2. Practical application in hazardous waste management.
3. Comprehensive use of GIS and mathematical modeling.

Weaknesses:

1. Limited scalability of the model to larger datasets. Enhance the scalability of the model to handle larger datasets and complex scenarios.
2. Focus on Mauritius may limit generalizability. Extend the study by applying the model to other countries or contexts to validate its versatility and applicability.
3. The sensitivity analysis could be more extensive. Include a comparative analysis with existing methods to highlight the advantages and limitations of the proposed algorithm. Conduct comprehensive sensitivity analysis to study the impact of varying parameters like facility costs and client distributions.

Discussions

Weakness 1.

The initial software used to develop the algorithm developed for this study had a limitation on instance input size, it was only reliably usable in cases where there were 7 Facilities and 125 clients. This meant that it opened the best out of 7 potential facility locations and assigned only 125 clients to these open facilities per calculation [2]. This limitation was with IBM ILOG CPLEX Optimization Studio 22.1.2 community edition Figure 5 [37]. This is a free version of the software. The commercial version of the software does not have the same size limitations. The IBM ILOG CPLEX Optimization Studio 22.1.2 Developer Edition was used to test and compare the performance of the algorithms in larger instances. This enabled the algorithm to solve larger instances of sizes, up 1000 facilities and 3000 clients.

Weakness 2

Appendix A [2] hosts the algorithm, this model is independent and separate from the data. A thorough analysis of the code in Appendix A will reveal that the algorithm can be used in all global contexts. This might mean renaming some of the variable names to suit the chosen area of application in whichever country might need to use the same model. This makes it usable globally. Appendix B [2][3] hosts a small more detailed demonstration to which a google earth map is used to visualize the median location. This is illustrated so that it is easier to adopt the same algorithm in global application. The problem in this case applied specifically to Mauritius a small island where no such research on ULABs had been conducted in over 30 years. However, the solution developed may be applied to other countries as well.

Weakness 3

A sensitivity analysis was performed on key parameters[2].

Parameters	Description	Value
		ESM-UFL
O	Number of Facilities	7
C	Maximum Number of clients	125
FI	Opening Cost MUR (Total or Shell 4100) (Indian oil or Engen 4400)	4100 4400
S= (x+y)	size of instance	132
x	*x* client locations	125
y	y facility locations	7

Figure 7. Small case sensitivity analysis [2][3]

The response variable AVG GAP (%) is the average gap of 30 different independent runs compared to the optimal solution obtained with CPELX Opl for the generated small case test problem [8]. The response variable AVG TIME is the average execution time for the 30 independent runs [8]. To allow a direct mapping and a clear visual comparison of the performance of the algorithms developed for this study compared to the [8] a total of 1440 trials were performed. The number of clients and the size of the instance affect the running time.

If an instance has many clients (C) or a size larger than 107 this affects the execution time. The AVG TIME shows that the number of open facilities (O) must be higher than three. The ANOVA results of the response AVG GAP (%) show that parameters C and Fi are the only statistically significant parameters. To find the best parameters for the algorithms, we focus on the response AVG GAP (%) [8]. The best parameter settings on O, C, Fi and S, according to AVG GAP (%), are 3, 50, 4100 and 107, for the ESM UFL respectively. C and Fi prolong the computation time.

2.3a Reviewer 4 comments [3]

The main issue, however, is whether such a model is appropriate or not. The k-median problem aims to minimize the distance from "customers" to "open facilities" usually a means to account for the transportation cost. It is hard to imagine why such an objective would be valuable here. The premise seems potentially that if there are facilities sufficiently close to demand regions or points, there would be more propensity to deliver these batteries to the collection facilities. Fair enough, but there are much better locational models for that purpose such as p-center as described again in [25] or maximal covering location problem (MCLP) of [28] or the maximum capture location problem (MCP) of [32] and one of its many generalization as in [26] or even a further but early refinement such as presented in [27]. If the premise, however, is to minimize some measure of transportation cost then the detailed account of transportation cost from the dispersed geography to these collection facilities as well from to the processing facilities.

Discussion

p-center is sensitive to outliers[3]. It is not as robust as the k-median. MCP and MCLP rely on estimates of demand, this being the first time this research is done in 30 years, all such estimates would be highly unrealistic. There is not enough data available to even envision a potential worst-case prediction. The transportation of hazardous waste is a whole different topic for future research, this is the first model which deals with the collection.

Considering that no research on the topic had been conducted in Mauritius for almost 30 years [5], there was not enough data that had been collected related to the area of research. So, the choice of the locational model was based on an attempt to make a realistic model from which future researchers would find meaningful, truthful and realistic information in terms of the data used. Thus, the authors chose a model that could make the best use of the available data. There is no data at all on the demand at each potential collection facility location because the research was pioneering.

If needed, the data would have to be an unreal prediction, while the focus was on delivering a realistic practical usable solution. The MCLP [28] was introduced to determine the locations of facilities of a fixed number that serve the most "potential demand" within the service standard [33]. As this is the first time since [5] the level of demand of ULABs, would be an extremely unrealistic wild guess at each potential facility location, considering the problem covers the whole country.

"To the best of our knowledge, existing works on facility location under Random Utility Maximization (RUM) all assume that the parameters of the RUM model are known with certainty and ignore any uncertainty associated with the estimates, with a tacit understanding that the parameters must be estimated in practice. Such an estimation can cause errors, and the decision-maker needs to cope with the fact that the "estimates of the choice parameters can significantly deviate from their true values". Ignoring such estimation errors would lead to bad decisions, as shown in several studies in the robust optimization literature" [34] The goal here is to maximize the worst case expected captured customer demand when the RUM parameters vary in the uncertainty sets [34].

There were only a few parameters that had enough data, which covered the whole of Mauritius, most types of data would be available in one city and not available in the next. Thus, the parameters that had available data were used to select a relevant locational model[2][3]. Human population data was available for every region, human population centroid data was available for every region, service station data for the whole country. There was no car density data for every region available, There was no data on which areas were light industrial areas, car density, car population distribution. Future work may be based on the data gathered after the implementation of this model.

The implementation and the model chosen was based on the availability of the actual true data at the central statistics office (CSO)and the national travel authority (NTA). There is no data available on the level of demand of the batteries so if another model or objective where to be chosen for example the maximal covering model which considers the weights of the demand and locations, it would not have been possible to adopt it in this case. The battery resellers were not willing to disclose data on their annual sales per region because this would disturb their daily business, so they would not give such information to their competition. There is no map which shows the density of cars in every region, just data on the importation and exportation of cars. There is no data which shows the light in the heavy industry of areas in Mauritius, just a shaded map which shows the regions of the population in every region. Thus, the lack of availability of data constraints the use of many other models. When it comes to solving economic decision problems that involve the accurate sitting of facilities to serve a given set of demands efficiently, the k-median objective function has been considered a better objective function than the mean or the p-center because of its robustness or its resistance to outliers [23]. The k-median algorithm is a variant of the k-means clustering technique. The aim of the k-means algorithm is to minimize the sum of squared distances between the data points and their respective cluster centroid. The data collected in many real-world applications may be contaminated with outliers, which can make traditional clustering methods such as K-means sensitive to their presence. Thus, robust clustering algorithms such as k-medians are used to produce more reliable outcomes [14].

2.3b Reviewer 4 comments [3]

Finally, I also like to mention that state of the art in solving p-median problems or the UFL problem has already progressed much farther than what is in the work (see for example,[30] and [31] for p-median and [29] for the UFL.)

Discussion

This is not true. The results of [2][3] can be compared against [29] [30] and [31]. The solutions in this case and the way of solving these problems can be seen in [11], they used the distance matrix approach which made it laborious and difficult to input the data or even solve the problem. All the tests on [2][3] prove it has a faster computation time than most of the leading solutions.

2.3c Reviewer 4 comments [3]

I also cannot see why in the implementation direct line distance is used, as ArcGIS or any web-based mapping application can find the exact distances from those demand centroids to the alternative facility locations using the actual road network.

Discussion

Euclidean distance between 2 points, direct line distance is the most common distance used with almost all modern Location science locational models. ArcGIS has built in features for location science that incorporate p-median and k-median models in all cases they use direct line distance. Arc GIS does not consider the opening cost of facilities in its calculation. Many uses and features of ArcGIS are not meant for optimum reduction of costs. They incorporate to many unnecessary features such as topography which makes it a tedious task when there are too many facilities and clients. The identification of clients and facilities is also extremely difficult even when the opening costs are not considered, it can be extremely difficult and time-consuming.

"These commercial tools rely on heuristics for solving location models, generally not providing technical details or solution quality assurance. This point has been noted by [38], who reported that ArcGIS did not derive the optimal solution for the LSCP instances. [39] also found that ArcGIS failed to optimally solve most of the 1059 MCLP problems examined. Moreover, the heuristics used by ArcGIS involve parameters that influence speed and solution quality, yet users are "prevented" from knowing or interacting with these parameters [40] [41]). Users are also forbidden from knowing or controlling the parameters of pre-processing steps in ArcGIS, such as assigning points to a network [42]. In practice, the lack of transparency and solution quality assurance raises issues of significance in reported findings." [34]

2.4a Reviewer 5 comments [2]

A model of reverse logistics that utilizes the ESMvere Cplex UFL problem to locate collection points for ULABs was simulated and designed to accommodate nineteen ULAB collection points. The nineteen coordinate points can be tested, they can be input on google maps to verify the accuracy of the solution visually. The location solutions were tested using ArcGIS and google maps. The solutions developed in this study are all usable in other industries and for other purposes besides the purpose of this study. This is because the models are separated from the data. A simulation solution has been developed that helps in the safe collection and disposal of ULABs. A network has been modeled that has the coordinate points of the nineteen optimum locations to establish the collection facilities for ULABs. I accept this very solid and important scientific paper.

Discussion

Thank you

3. CONCLUSION

The contributions of [1][2][3] are theoretical developments, computational algorithms and practical applications. The techniques used to solve [2][3] are exact Integer Binary heuristic solution techniques, that can be applied in many other contexts. The methodology eliminates the use of the cost distance matrix and replaces it with the Euclidean distance formular resulting in a realistically usable algorithm proven by the use of actual GPS coordinates. This is the first methodology to achieve this when solving the UFL problem. The software code is available for every academic to try investigate and test for themselves and come up with their own analysis.

Acknowledgements

I would like to thank my parents, My Mum who has made all my research efforts possible., My whole family, my darling, my colleagues. The people of Zimbabwe, the people of Mauritius, the people of Africa, and all the people of the world.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

4. REFERENCES

1. Mvere, Emmanuel (2025) An Analysis of the Used Lead Acid Battery (ULAB) Collection Network Mauritius Using the ESMVERE- R-Hamming-k-median Clustering Method., *International Journal of Science Academic Research*, Vol. 06 No. 1, 2025, pp. 9113-9126.
2. Mvere, E.S, Callychurn, D.S & Hurreeram, D. K (2024) The Near Optimal Siting of Hazardous Waste (Used Lead Acid Battery (ULAB)) Collection Facilities in the Republic of Mauritius Using the ESMVERE CPLEX UFL Problem

Solver Algorithm, *American Journal of Operational Research*, Vol. 14 No. 1, 2024, pp. 12-27. doi: 10.5923/j.ajor.20241401.02.

3. Mvere, E.S, Callychurn, D.S & Hurreeram, D. K (2025). The Near Optimal Siting of Hazardous Waste (Used Lead Acid Battery (ULAB)) Collection Facilities in the Republic of Mauritius Using the ESMVERE CPLEX k-median Problem Solver Algorithm, *American Journal of Operational Research,* Vol. 15 No 1 pp. 12-27. doi: 10.5923/j.ajor.20251501.01 02.
4. Voynick, S. (2011). Trade and environment a teaching case: The Basel ban and batteries. *Institute for trade and commercial diplomacy*.
5. Phoolchund, H. (1995). Occupational and environmental health in Mauritius: a review of trends and recent studies. *Elsevier Science Ltd* pp. pp.251–255.
6. Jain, K. & Vazirani, V. (1999). Primal-dual approximation algorithms for metric facility location and k-median problems. *in* 'Proceedings of the 40th Annual Symposium on Foundations of Computer Science'. FOCS '99. p. 2.
7. Li, S.(2013) A 1.488 approximation algorithm for the Uncapacitated facility location problem, *Information a nd Computation*, *Vo* 222, 2013, Pages 45-58, doi.org/10.1016/j.ic.2012.01.007.
8. Peidro, D., Martin, X., Panadero, & Juan, A. (2024). Solving the Uncapacitated facility location problem under uncertainty: a hybrid tabu search with path-relinking simheuristic approach
9. Cacioppi, P.& Watson, M. (2014). *A Deep Dive into Strategic Network Design Programming: OPL CPLEX Edition*. Northwestern University Press.
10. Mistry, S (2019) Optimization-Facility-Location [source code]. https://github.com/samarthmistry/Optimization-Facility-Location/blob/master/UFL.cpp
11. Cornuejols, G., Nemhauser, G & Wolsey, L (1984). The Uncapacitated Facility Location Problem. Cornell University
12. Balinski, M. (1966). On finding integer solutions to linear programs In: *Proceedings of the IBM Scientific Computing Symposium on Combinatorial Problems*, pp. 225-248.
13. Srinivasan, G. (2012). Mod-08 Lec-32 Location problems -- p median problem, Fixed charge problem. *https://www.youtube.com/watch?v=82OUSjOM2bg*
14. Baggioni, A. G., and Surendran, S.(2024) A penalized criterion for selecting the number of clusters for K-medians Journal of Computational and Graphical Statistics, 2024•Taylor & Francis
15. Moshkovitz, M., Dasgupta, S., Rashtchian, C. and Frost, N., (2020), November. Explainable k-means and k-medians clustering. In *International conference on machine learning* (pp. 7055-7065). PMLR.
16. Manish Suyal, Sanjay Sharma (2024). A Review on Analysis of K-Means Clustering Machine Learning Algorithm based on Unsupervised Learning. *Journal of Artificial Intelligence and Systems,* 6, 8–95. https://doi.org/10.33969/AIS.2024060106.
17. Anand, A. and Lee, E., (2024) Separating k-Median from the Supplier Version University of Michigan, Ann Arbor *arXiv preprint arXiv:2401.13819v[csDS]24 Jan 2024*.
18. Whelan, C., Harrell, G. & Wang, J. (2015). Understanding the k-medians problem. *International Conference Scientific Computing* pp. pp.219–222.
19. Olea, J.L., Prallon, B., Qiu, C., Stoye, J., and Sun, Y. (2024) Externally Valid Selection of Experimental Sites via the k-Median Problem
20. Glover, F., Hanafi, S., GuemrI, O. & Crevits I,(2018) A simple multi-wave algorithm for the Uncapacitated facility location problem
21. Jain, K. & Vazirani, V. (2001). Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. *Journal of the ACM - JACM*.
22. Jain, K., Mahdian, M. & Saberi, A. (2002). A new greedy approach for facility location problems. *in* 'Proceedings of the thirty-fourth
23. Herold, M.G., Kipouridis, E. and Spoerhase, J., (2024). Clustering to Minimize Cluster-Aware Norm Objectives. arXiv preprint arXiv:2410.24104.
24. Pradeepmon, T.G., Paul, B.,(2012). A simulated annealing-based hybrid algorithm for solving UFL problems. *International Journal of Applied Management Science*, Vol.4.No.3,pp.274-283
25. Hakimi, S.L., 1964. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research,* 12(3), pp.450-459.
26. Berman, O. and Krass, D., (2002). The generalized maximal covering location problem. *Computers & Operations Research,* 29(6), pp.563-581.
27. Berman, O., Krass, D. and Drezner, Z., (2003). The gradual covering decay location problem on a network. *European Journal of Operational Research,* 151(3), pp.474-480.
28. Church, R. and ReVelle, C., (1974), December. The maximal covering location problem. *In Papers of the Regional Science Association* (Vol. 32, No. 1, pp. 101-118). Berlin/Heidelberg: Springer- Verlag.
29. Erlenkotter, D., (1978). A dual-based procedure for Uncapacitated facility location. *Operations Research,* 26(6), pp.992-1009.
30. Hansen, P. and Mladenovic, N., 1997. Variable neighborhood search for the p-median. *Location Science,* 5(4), pp.207-226.
31. Reese, J., (2006). Solution methods for the p-median problem: An annotated bibliography. *NETWORKS: An International Journal,* 48(3), pp.125-142.
32. ReVelle, C., (1986). The maximum capture of "sphere of influence" location problem: Hotelling revisited on a network. *Joumol of Regional 5cience,* 26(2), pp.343-358.
33. Chen, H., Murray, A.T. & Jiang, R. Open-source approaches for location cover models: capabilities and efficiency. *J Geogr Syst* 23, 361–380 (2021). https://doi.org/10.1007/s10109-021-00350-w
34. Dam, T., Ta, A.T., and Mai, T.(2003) Robust maximum capture facility location under random utility maximization models, *European Journal of Operational Research,* Vol 310, Issue 3, pp 1128-1150,
35. Tong D., Murray, A.T.(2009) Maximising coverage of spatial demand for service. Pap Reg Sci 88(1):85– 97. https://doi.org/10.1111/j.1435-5957.2008.00168.x
36. Tong D, Wei, R. (2017) Regional coverage maximization: alternative geographical space abstraction and modeling. *Geogr Anal* 49(2):125–142. https://doi.org/10.1111/gean.12
37. IBM (2025) Decision Optimization at a glance. https://www.g2.com/products/ibm-decision-optimization/pricing *pp. 2025–02–25*
38. Woodhouse, S., Lovett, A., Dolman, P., Fuller R (2000) Using a GIS to select priority areas for conservation. Comput Environ Urban Syst 24(2):79–93. https://doi.org/10.1016/S0198-9715(99)00046-0

39. Murray, A.T., Xu, J., Wang, Z., Church, R.L. (2019) Commercial GIS location analytics: capabilities and performance. Int J Geogr Inf Sci 33(5):1106–1130. https://doi.org/10.1080/13658816.2019.1572898
40. Church R.L. (2002) Geographical information systems and location science. Comput Oper Res 29(6):541– 562. https://doi.org/10.1016/S0305-0548(99)00104-5
41. Murray A.T. (2010) Advances in location modeling: GIS linkages and contributions. J Geogr Syst 12(3):335–354. https://doi.org/10.1007/s10109-009-0105-9
42. Gaboardi, J.D., Folch, D.C., and Horner, M.W. (2020) Connecting points to spatial networks: effects on discrete optimization models. Geogr Anal 52(2):299–322. https://doi.org/10.1111/gean.12211
43. Ahmadian, S. (2017). Approximation Algorithms for Clustering and Facility Location Problems. PhD work. University of Waterloo.

APPENDICES

Appendix A

The ESMvere CPLEX UFL Problem Solver Model

The ESMvere CPLEX UFL problem solver algorithm is a software invention that solves the UFL problem. It is a tool that aids, and guides management in making sound decisions on where to locate facilities. It aids management in the decision of where the best places are, lowest cost places, the minimum distance places to locate facilities considering that there are opening costs and the Euclidean distances, direct line distances of GPS coordinate points?

The CPLEX Opl Code

```
int n =...;                    //number of facilities
int m =...;                    //number of clients declared in data files
        range ulabfacilities = 1..n;
        range ulabcities = 1..m;
        int fi[ulabfacilities] =...;
tuple Location {float x; float y; }
tuple edge {int i; int j; }
setof (edge) edges = {< i, >
j | i in ulabfacilities, j in ulabcities };//defines an edge
float c[edges]; //An Array of each edge on the plane measured separately
Location ulabcityLocation [ulabcities];
Location ulabfacilityLocation [ulabfacilities];
execute {
function getDistance (ulabcity1, ulabfacilitycity1) {        //Euclidean distance between each facility and client
return opl.sqrt(opl.pow(ulabcity1.x − ulabfacility1.x, 2)
        + opl.pow(ulabcity1.y
−ulabfacility1.y, 2)) }
for(var i inulabfacilities) {
ulabfacilityLocation[i].x;
ulabfacilityLocation[i].y;
}
for(var j inulabcities) {
ulabcityLocation[j].x;
ulabcityLocation[j].y;
}
for(var e in edges){ //distance between each edge"heuristic"
c[e] = getDistance(facilityLocation[e.i], cityLocation[e.j])
} }
dexpr float OpeningCost = sum(i in ulabfacilities) fi[i] ∗ y[i];

dexpr float TotalDistance = sum(e in edges) c[e] ∗ x[e];
Minimize OpeningCost + TotalDistance;        //objective function
subject to {                   //constraints from p-median[9]
forall(j in ulabcities: j >= 1)
sum ( i in ulabfacilities) x[< i, j >] == 1;
forall(i in ulabfacilities , j in ulabcities)
x[< i, j >] <= y[i];
sum(i in ulabfacilities)
        }
```

INPUT FILE

```
Ex 3.dat

/*****************************************
 * OPL 20.1.0.0 Data
 * Author: Emmanuel Siyawo Mvere
 * Creation Date: OCT 28, 2022 at 6:58: PM
 *****************************************/

n=7;

m=21;

ulabcityLocation=[<0.1 6.4> <0.2 0.9> <0.75 2.1> <1 6.9> <2.44 5.6> <3.1 3.4>
  <3.45 0.9> <4.25 5.6> <5.15 2.2> <5.6 0.95> <6.1 6.95> <7.55 6.9> <7.9 4.4>
  <9.2 1> <9.35 3.5> <9.65 4.4><10.55  5.7><12.3 2.25> <12.4 4.45> <12.55 7>
  <12.65 1>];

ulabfacilityLocation=[<0.7 3.8> <2.2 1.2> <4.3 7.3> <5.8 4.75> <7.3 1.3> <10.2
  2.5><10.8 7.3> ];

fi=[30, 20, 50, 20, 40, 60, 10];
```

Figure 8. Input file

In this case the input are x and y coordinates [<x y><x y>] can be replaced by GPS Coordinates [< -20.172544 57.4883317 ><-20.165595 57.492625>] compare with a distance matrix of the same coordinate points.

APPENDIX B

[10] Mistry, S (2019) Optimization-Facility-Location [source code]. https://github.com/samarthmistry/Optimiza... Location/blob/master/UFL.cpp

```
// Include the header file
#include<ilcplex/ilocplex.h>
#include<vector>
// Use the macro before class definition
ILOSTLBEGIN
// Define two-dimensional variable array
typedef IloArray<IloNumVarArray>IloNumVarArray2;
// Defining Solve function having IP formulation
int main()
{
        // Create the environment object
        IloEnvenv;
        // Write try-except clause for exception handling
        try
        {
                // Define input file and load values
                const char* filename = "UFL.dat";
                ifstream file(filename);
                if (!file) {
                        cerr<< "No such file: " << filename <<endl;
                        // Define output file
                        ofstreammyfile("UFL_output.txt", ios_base::app);
                        throw(-1);
                }
                // Load data
// First array is client demand, second array is distribution cost matrix (between a facility and a client), and last array is facility fixed cost
                IloNumArrayclient_dem(env), dist_cost(env), fac_cost(env);
                file >>client_dem>>dist_cost>>fac_cost;
```

****