

Document downloaded from:

<http://hdl.handle.net/10251/168475>

This paper must be cited as:

Palanca Cámara, J.; Jordán, J.; Bajo, J.; Botti Navarro, VJ. (2020). An energy-aware algorithm for electric vehicle infrastructures in smart cities. *Future Generation Computer Systems*. 108:454-466. <https://doi.org/10.1016/j.future.2020.03.001>



The final publication is available at

<https://doi.org/10.1016/j.future.2020.03.001>

Copyright Elsevier

Additional Information

An Energy-aware Algorithm for Electric Vehicle Infrastructures in Smart Cities

Javi Palanca^a, Jaume Jordán^a, Javier Bajo^b, Vicent Botti^a

^a*Institut Valencià d'Investigació en Intel·ligència Artificial (VRAIN), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain*

^b*Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Campus Montegancedo, Boadilla del Monte, Madrid 28660, Spain*

Abstract

The deployment of a charging infrastructure to cover the increasing demand of electric vehicles (EVs) has become a crucial problem in smart cities. Additionally, the penetration of the EV will increase once the users can have enough charging stations. In this work, we tackle the problem of locating a set of charging stations in a smart city considering heterogeneous data sources such as open data city portals, geo-located social network data, and energy transformer substations. We use a multi-objective genetic algorithm to optimize the charging station locations by maximizing the utility and minimizing the cost. Our proposal is validated through a case study and several experimental results.

Keywords: electric vehicle, charging station, genetic algorithm, energy, smart city, multi-objective, evolutionary algorithm, deap, Peru

1. Introduction

An increasing number of people are talking about the electric vehicle (EV), but there are many who still have not decided to buy an electric car, not only because of its cost, but also because of the lack of confidence in the network being able to withstand this new demand and the absence of public charging stations. More than 900 million vehicles circulate on the planet and three million of them are already electric. China dominates the demand with 580 thousand units sold and the United States is second with 280 thousand units. In order for the market to develop faster, electric companies must begin to plan the installation of infrastructure in homes and highways in order to provide users with charging stations. So far, the most logical thing is to install them in supermarkets, shopping centers, and parking lots.

However, in countries such as Peru, the introduction of the EV has been slowed down by the lack of regulation and especially by the absence of infrastructure. This lack of a public EV charging infrastructure is partly due to the complexity of properly structuring the deployment of these stations. Creating

this network of charging stations is expensive and complex and there is a concern on the part of local authorities to appropriately select locations for the installation of these infrastructures.

This paper presents a proposal to solve the problem of choosing the most suitable configuration to create an EV charging network in a city. In order to do this, it is necessary to decide the location of the charging stations to be installed among all of the possible locations in the city and the number of charging poles that each station will have (which will be different for each station depending on the need). Given the large number of possible solutions, the search space is large enough to propose the use of a genetic algorithm.

The genetic algorithm presented here has been adapted to manage the information extracted from a smart city and to take advantage of it. The results of this algorithm are improved thanks to the use of population distribution information, traffic intensity information by zones, activity in social networks, location of electrical transformer substations, etc.

The contents of the paper are organized as follows. Section 2 reviews existing platforms and tools about EV charging stations as well as related work about locating EV charging infrastructures using different algorithms. Then, Section 3 formalizes the problem that we aim to solve and the notation that we follow. The multi-objective evolutionary algorithm that solves the proposed problem is presented in Section 4. Section 5 presents a case study in the city of Lima (Peru) with some experimental results. Finally, we present final conclusions and future work in Section 6.

2. Related work

Currently, there is significant effort being made in the consolidation of the Electric Vehicle (EV) in society. The supply of different car manufactures regarding EV and hybrid vehicles has been increasing over time. Furthermore, governments, administrations, and town councils are also providing funds and opportunities to establish different EV charging stations.

The users of EV have different tools for finding existing EV charging stations and can even plan their route. For instance, there are several online tools such as those in [1, 9, 3] that provide the location of alternative fueling stations for the United States and Canada. Additionally, there is a more complete tool in [5], which is a mobile application and an online tool for finding EV charging stations worldwide that includes information about the type of connectors. Similarly, there are several alternative tools such as those in [26, 4]. Finally, in the well-known Google maps application [15], a simple search for “EV charging stations” will provide the charging stations that are in the area shown on the map.

In the field of EV charging, there is a significant research work that focuses on optimizing the charging of a set of EVs in a decentralized way [14]. Other works try to optimize the use of the electric supply network by reducing the impact of the EV on it [22].

There are also several guidelines such as those in [10, 2] to help administrations and governments implement EV charging stations and to determine where

to place them. Similarly, the guide in [27] is directly based on the experience of governments in EV planning and implementation processes.

An important problem regarding EV charging stations is to determine their locations. It is critical to provide an EV charging infrastructure that is sufficient to cover demand in order to promote its use. This problem has been addressed in different ways, as we analyze below in the following paragraphs.

First of all, there still exists a limitation in the adoption of the EV because of a phenomenon called “range anxiety”. This phenomenon is due to a charging station infrastructure that is insufficient, which makes users unsure if they will be able to cover the desired distance within their EV range without having options to charge their batteries [23]. Thus, there are also several studies that analyze the charging behaviours of users [13, 12].

Mathematical programming is used in [28] to determine the best places to establish alternative fuel stations in order to maximize the number of vehicles served. Another work uses a conceptual optimization model that focuses on a corridor in which the goal is to choose the number of stations and the charging power at each station in order to provide enough service while minimizing the social cost [24]. There is also a tool in [20] that focuses mainly on the voltage analysis needed at charging stations to evaluate the effects of typical loads in the charging poles and the distribution network.

The work in [32] uses driver behaviour, vehicle performance, battery attributes, travel profiles, environmental conditions and charging infrastructure to strategically locate the charging stations in order to maximize the performance of the use of EVs. A similar approach presented in [30] tries to maximize the service of electric taxis and their charging by using a genetic algorithm that takes into account the range that a taxi can travel, its charging time, and the capacity of the charging stations. Another work in [8] also uses a genetic algorithm to select the location and type of charging poles, using a grid in which each cell contains the number of trips that end in it. Then, the most popular destinations are selected as the potential locations for charging stations. Similarly, the model in [21] solves a maximum demand coverage problem to locate the EV charging stations in a weighted network in which the weight of the arcs is the number of cars from the origin to a destination.

An agent-based, decision-support system that identifies patterns in driving activities and residential EV ownership is presented in [29]. This data is used to deploy a charging infrastructure. In a similar way, He et al. [16] mainly consider EV driving ranges to optimize the location of charging stations so as to maximize the path flows by using a heuristic algorithm. Alternatively, the users’ travel destinations are considered as Points of Interest (PoIs) in [31]. The model proposed in the cited work ranks the PoIs according to their attraction for EV users, and two approaches are considered. One approach uses a method to obtain the maximum demand coverage to optimize the location of charging points. The second approach uses an iterative method to penalize the PoIs that are close to existing charging points.

Finally, none of the cited works mentioned above combine all of the data that can be extracted from a city, i.e., population, traffic or frequent routes,

social data, available energy or distribution network locations, and the cost of each charging pole and station.

Nevertheless, our previous works of the current proposal presented in [18, 17] do use a genetic algorithm to obtain near-optimal solutions by integrating a set of heterogeneous sources of data. These proposals obtain data from open data web portals for the population and traffic of the studied city, data from social networks to measure the geo-located activity of the users, etc. Hence, the goal of these proposals is to deploy an EV charging infrastructure by using a genetic algorithm to locate charging stations by starting from a set of POIs and using all of the data obtained from the city.

Thus, the main goal of our new proposal is to improve the previous ones by gathering more crucial data such as energy and using better optimization techniques to locate EV charging stations in a smart city. Additionally, other minor improvements are also considered. The main novelties of the current proposal are the following:

- The introduction of a multi-objective genetic algorithm to separate the parameters that measure utility and cost instead of considering them in a single fitness function. The solutions obtained are also in the Pareto frontier to guarantee the optimization of all the objectives.
- The inclusion of power transformer substations of the city to take into account the distance to connect the new charging stations to be locate. This is a crucial aspect in order to determine whether or not place a charging station in an area, considering the availability of energy infrastructure.
- Improvement in the treatment of geo-located social network data by the application of a geohashing algorithm that concentrates the individual messages into zones. These zones are then tagged with the number of social network messages generated in the area. This reduces the computation significantly in environments with a huge amount of data that is extracted from social networks.
- A case of study and experimental evaluation in the city of Lima (Peru), from which we gathered all the necessary data. This involves the challenge of obtaining the new data of a city as well as the challenge of testing our system in another context.

3. Problem Definition

This work aims to find a solution to the problem of locating a set of Electric Vehicle (EV) charging points in a city. Cities are large and complex systems, and, thanks to smart city technologies and open data, we have a lot of information (both static and real-time) that can be used to try to select the best places to install such stations.

We rate a solution as good if it takes into account variables such as the amount of population that a station may serve, the amount of traffic of roads

near a station, or whether the area where the station is located has activity during the day (commercial and service districts) or receives its main activity at the end of the day (residential districts).

There are more infrastructures in a city that may be useful as inputs for this problem, some of which can make the installation of charging stations cheaper: being close to any of the energy transformer substations found in cities, being connected to a center that has enough power to provide the service of charging EVs, or being connected to a center that reduces the public works required to bring energy to the selected point. Those are discussed further in the following sections.

Since evaluating every coordinate of a city in order to determine where to place a set of EV charging stations would be too big of a problem, in this work, we use a genetic algorithm that decides on a sub-set of pre-defined points that an expert has properly selected. This set of points (I) is made of Points of Interest and is defined as:

$$I = \langle (x_0, y_0), (x_1, y_1), \dots, (x_n, y_n) \rangle$$

where each pair of (x_i, y_i) are the coordinates of a Point of Interest. A Point of Interest (PoI) is selected in this problem because it is a candidate for being a future EV charging station location. These points are usually in places where it is interesting to place such stations. For example, they are close to busy areas; there are parking spaces available where you can stop for a long enough period of time to charge the vehicle; or they meet other criteria chosen by the expert. However, this set of points is usually too large, and since it is not possible to deploy as many charging stations in the city, an algorithm such as the one proposed here is needed to determine which of all these points are more optimal to install EV charging stations.

To make this decision, it is important to feed the algorithm with useful data that helps it to optimize the solution. This data may be taken from different sources; these are usually open data platforms that the city makes available to citizens. However, it is also common to obtain more specialized information directly from the technicians of the city council or the electric companies that provide services in the city. For this problem, we have selected the following data:

1. **Cadastral information:** Knowing how many people live in each zone of the city may be useful in determining where the demand of charging stations may be higher. That is why we use the population information (P) as:

$$P = \langle (C_1, p_1), (C_2, p_2), \dots, (C_n, p_n) \rangle$$

where C_i is a polygon that represents a cadastral block in the city with population data (p_i). A polygon C_i is made of a set of coordinates, where the first and the last point are the same since it is a closed polygon:

$$C_i = \langle (x_{i_1}, y_{i_1}), (x_{i_2}, y_{i_2}), \dots, (x_{i_1}, y_{i_1}) \rangle$$

2. **Traffic information:** The number of vehicles circulating in a given area is also a good indicator for making decisions about the need to install a station near high traffic areas. The traffic information (T) is defined as:

$$T = \langle (L_1, t_1), (L_2, t_2), \dots, (L_n, t_n) \rangle$$

where L_i is a polyline that defines a street segment with a measure of traffic t_i :

$$L_i = \langle (x_{i_1}, y_{i_1}), (x_{i_2}, y_{i_2}), \dots, (x_{i_n}, y_{i_n}) \rangle$$

3. **Social Activity:** Another source of useful information for determining if a certain area of the city is active is monitoring social networks in real time. By capturing geo-located statuses from social networks, we can determine with great precision where a sufficiently representative percentage of the population is spending their time [19]. In our case we used tweets from the social network Twitter. This metric gives us very interesting information, since knowing where people spend the day is even more useful than the population information because it gives us some clues about their mobility. To represent the Twitter geo-located activity in a city, we use tweets with geo-location information, which is a latitude-longitude pair. However, managing every coordinate of a city introduces a high computational cost. This is due to the high granularization of the coordinate values (due to the fact that they are represented with floating numbers). Therefore, we encode them to a simpler format called *geohash* [25]. This is a public domain geocoding system that allows us to put coordinates in buckets of grid shape. These buckets are encoded as a hash made of a short string of numbers and letters. The number of characters represents the precision of the bucket. The longer it is, the more accurate it is and also the smaller the area covered by the bucket. Thanks to this encoding system we can represent the social activity in a city as:

$$A = \langle (h_1, a_1), (h_2, a_2), \dots, (h_n, a_n) \rangle$$

where h_i is a geohash code with 7 characters of precision (which is approximately 150 meters of precision) and a_i is the quantity of social activity (in our case, tweets) that have fallen inside the bucket h_i . Note that we are using tweets in this example, but you can impute any type of geo-localized data to a bucket as long as this data represents some kind of social activity in the city (e.g., cell phone calls). In this work, we have used social information from Twitter because it has the following characteristics: it is open and accessible data, available through a public API for anyone; it is a very active social network that generates 500 million tweets every day around the world; approximately 1% of the tweets published are geo-localized. However, it should be noted that the number of tweets collected is not a determining value, since what we take into account in this work is the distribution of these tweets throughout the geography to be analyzed.

Even so, as the algorithm interprets this data as social activity, it can always be improved by integrating data from other sources such as other social networks or cell phone calls if this information is available.

4. **Transformer Substations:** Since EV charging stations have a high requirement for energy consumption, arranging them in appropriate places to minimize the cost of the work of laying electricity and ensure that they will have sufficient available power is a good idea. In the algorithm proposed in this work, we try to bring these infrastructures as close as possible to the transformer substations that are distributed throughout the city. A transformer substation is an electrical installation that receives high voltage energy (30 kilovolts) or medium voltage energy (10, 15 or 20 kilovolts) and delivers medium or low voltage for use by end users, normally at 400 volts in three-phased systems and 230 in single-phased systems. Since we are using the location and properties of these substations we define:

$$E = \langle (x_1, y_1, e_1), (x_2, y_2, e_2), \dots, (x_n, y_n, e_n) \rangle$$

where x_i and y_i represent the coordinates of a substation (latitude and longitude) and e_i is the available energy at the substation. This value of available energy helps the algorithm to determine how many charging poles the charging station may have and what power each one can have.

When available, all of this data provides very useful information to make the decision of where to put the charging stations. The set of Points of Interest is a required pre-requisite since we are going to make the decision based on an initial set of possibilities. This set I can be as large as we want because the input to the algorithm is the number of charging poles that we plan to put in the city (usually each country has its own recommendation on the proportion of EVs to charging stations). The rest of the data introduces useful information that helps the algorithm to make an informed decision. An example visualization of this data in the San Isidro district of Lima (Peru) is shown in Figure 1. However, there are some constraints that may limit our decision process. These constraints force us to be adaptive, as is shown in next subsection.

3.1. An adaptive algorithm for environments with constraints

When we propose how to build a solution for the placement of EV charging stations, we need to take into account different inputs from a city. All of these inputs help the evolutionary algorithm that we present in the next section to optimize the solution. However, a city is a very complex system that cannot be treated as a static object. A city grows year after year, incorporating new elements and even eliminating some of the existing ones, and the algorithm must be adaptive to manage these kinds of constraints in the environment, depending on when it is run. One element that forces the algorithm to be adaptive is the possibility of previously existing charging stations in the city. These stations may be considered as a constraint to the algorithm since they should not be ignored or disabled. However, they could see their capacity increased as part of

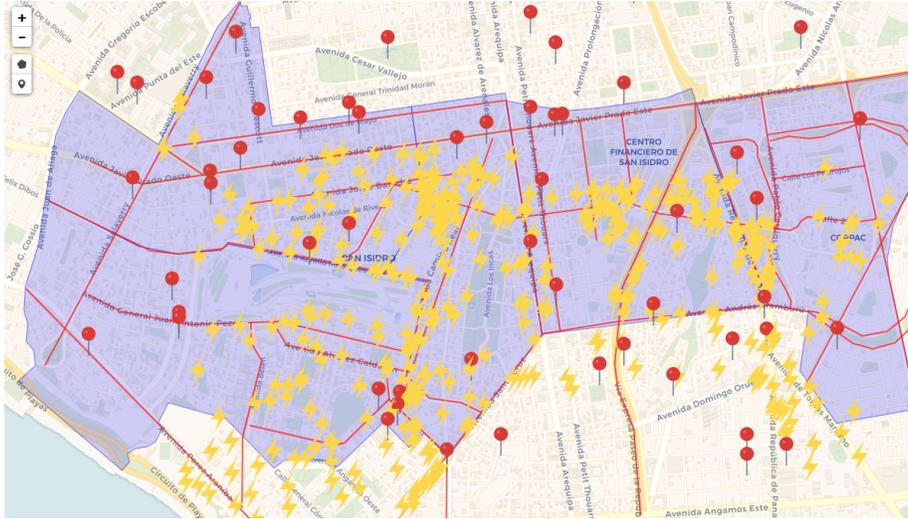


Figure 1: Example of input data for the San Isidro district of Lima (Peru). Legend: Pins represent Points of Interest. Lightning bolts represent Transformer Substations. Blue polygons depict cadastral information blocks and red lines depict traffic information.

the solution proposed by the algorithm (i.e., doubling the number of charging poles). That is why stations of this type that are already installed in the city should be viewed as a special case, where their current capacity is a restriction that cannot be diminished but may be increased.

This constraint is introduced as an input to the algorithm as:

$$F = \langle f_0, f_1, \dots, f_n \rangle$$

where $f_i \in \mathbb{N}_0$ is the number of previously installed stations in the Point of Interest $I_i \in I$. When f_i is 0, there is not a constraint for PoI I_i .

The adaptability of the algorithm to this restriction is presented in this work, where we manage it as a special case that determines the feasibility of an individual.

4. Multi-objective Evolutionary Algorithm

The nature of the problem we aim to solve entails using multi-objective optimization algorithms. We need to maximize the utility of locating a set of EV charging points while also minimizing the cost of this infrastructure. These are the two objectives to be optimized in the problem. Furthermore, the complexity of this problem makes other techniques inefficient considering the wide range of possibilities of locating stations in a considerably large number of positions and also the different number of charging poles in each station. We also have to note that each station located in a specific point has an area of influence that covers the population, traffic, and social network activity. Similarly, the cost

depends on the number of charging poles and the distance to the energy points. Therefore, it is a hard problem to solve due to the number of different variables to manage as well as the geometric calculations that are involved regarding the areas of influence and intersections in order to take into account the coverage of each station.

Of the different genetic algorithms that support multi-objective optimization, we used the well-known Non-dominated Sorting Genetic Algorithm II (NSGA-II) [7]. The main advantage that we exploit of NSGA-II with respect to a genetic algorithm with a single fitness function is the multi-objective optimization that obtains solutions near the Pareto frontier. This means that the algorithm attempts to obtain near-optimal solutions, by considering all of the objectives, and, hence, for any solution, one objective value cannot be improved without decreasing another objective value (from known solutions). In other words, the solutions returned by the NSGA-II algorithm are not dominated.

We use the $\mu + \lambda$ evolutionary algorithm implementation of the *deap*¹ library of *Python*. The μ value represents the number of individuals to select for the next generation, which we consider the population size of our problem. Then, the λ value is the number of children to produce at each generation, which we have defined as half of the population size in order to generate a significant number of children and to select the best among them.

The NSGA-II algorithm uses specific genetic operators for selection, crossover, and mutation. For the selection operator, differently from a general genetic algorithm, the NSGA-II uses its own selection operator that chooses the best individuals with respect to the Pareto frontier. The crossover operator that we use in our implementation is the two-point technique, in which two points of the chromosomes are selected and everything between those two points is swapped between the parents. Finally, we use the uniform integer technique as a mutation operator, which generates a new integer value within a provided range with an independent probability of each attribute being mutated.

4.1. Energy-aware Infrastructure Location

One of the main considerations in placing an EV charging station is its distance to an available electric transformer substation. Depending on the city council and the power company involved, there will be a specific limit distance to which it is feasible to build a connection. Also, within this limit, there is an associated cost of installing the electric power line from the transformer substation to the new charging station, for simplicity, in this case, we consider the cost to be linear with respect to distance. Additionally, there should be enough free power capacity in the transformer substation to supply the charging station. However, in this work, we assume that there is enough free power capacity in the transformer substations that we introduce in the algorithm.

Therefore, before starting the evolutionary algorithm, a pre-calculation can be done to find the distance to the nearest transformer substation for each PoI.

¹<https://github.com/DEAP/deap>

Then, during the execution of the evolutionary algorithm, the only calculation regarding the energy cost to be done is the evaluation of the fitness of each individual. For each PoI where a charging station is going to be placed, we compute the cost of connecting it to the nearest electric transformer substation (Θ_E):

$$\Theta_E = \begin{cases} base_cost \cdot distance, & \text{if } distance \leq limit_distance \cdot 1.05 \\ (base_cost \cdot distance) \cdot 2, & \text{otherwise} \end{cases}$$

where *base_cost* is a constant that is specified as a configuration parameter; and *limit_distance* is the maximum distance that the city council or power company allows to make a connection from a transformer substation to a charging station. Therefore, the cost of connecting energy to a charging station is linear with the distance, while this distance is lower than the specified limit (plus a 5% margin); then the cost doubles with the distance.

4.2. Fitness Function

The attributes that are considered to determine the activity around a PoI are population, traffic, and published tweets. The population can be specified as blocks corresponding to sets of houses. The traffic should be specified by streets; the more detail, the more accurate the algorithm can be. The published tweets must be geo-located so that they can be easily included in the corresponding PoI area. Additionally, the geo-located tweets are preprocessed using a geohashing technique in order to reduce computation. Hence, a collection of millions of tweets in a city is divided into small areas noting the number of tweets in each small area (similarly to the population blocks).

To evaluate each feasible individual, a Voronoi diagram is generated taking into account only the active PoIs, that is, the ones in which a station is going to be placed with a specific number of charging poles greater than zero. Then, the attributes (i.e., population, traffic, and tweets) to consider inside the influence area of the PoI must be inside the corresponding Voronoi polygon and inside the influence radius specified in the configuration of the problem. Hence, we define the **influence area** of a PoI as the zone of the geometric intersection of the Voronoi polygon with the circumference centered in the PoI with a radius given as input parameter of the problem (normally, it will take values of 50-1000 meters, depending on the other supplied input data). The reason for considering both the Voronoi polygons and the circumference around the PoI is to make the influence area more accurate in the case of near PoIs in which the circumferences collide. Hence, the intersection of the Voronoi polygon and the circumference impedes considering the same area by two or more PoIs at the same time. An example of the influence area of the PoIs is shown below in Figure 4 of Section 5.

We use the following notation: Φ is an individual, and each $cp_i \in \Phi$ is the number of charging poles (between 0 and a maximum determined by the configuration of the problem) to be placed in a station located at the PoI.

The utility of an individual Φ is defined as the sum of the attributes that cover the area of the corresponding PoI (i.e., population, traffic, and social activity). This is done for every PoI cp_i in which at least one charging pole is placed. The equation is as follows:

$$utility = \sum_{\forall cp_i > 0 \in \Phi} (\omega_P \cdot P_i + \omega_T \cdot T_i + \omega_A \cdot A_i) \quad (1)$$

where P_i is the covered population of the station obtained by the geometric intersection of the Voronoi polygon of the PoI, the polygons corresponding to population blocks and the influence area; T_i represents the covered traffic of the streets in the influence area of the PoI; A_i is the number of tweets that are geo-located in the influence area of the PoI; and each ω is the corresponding weight of the different attributes that are considered for the utility.

The cost (Θ) of an individual Φ is defined as:

$$\Theta = \sum_{\forall cp_i > 0 \in \Phi} (\Theta_{cp} \cdot cp_i + \Theta_{E_i}) + \Theta_S \cdot |\Phi| \quad (2)$$

where Θ_S is the fixed cost of a station, Θ_{cp} is the fixed cost of a charging pole, cp_i is the number of charging poles at each PoI of the individual, Θ_{E_i} is the energy cost of cp_i , and $|\Phi|$ is the number of activated stations in Φ where at least one charging pole is located (and hence a station is placed). In other words, $|\Phi|$ is the number of $cp_i \in \mathbb{N}_1$ (where $\mathbb{N}_1 = \{x \in \mathbb{N} | x > 0\}$).

Finally, the **fitness** function involves the **maximization of the utility** and the **minimization of the cost** to obtain solutions that comply with the Pareto frontier.

4.3. Feasibility and Penalty Function

In order to achieve solutions that respect the constraints, we defined specific feasibility functions to check whether or not an individual is feasible. Additionally, these functions allow us to apply a penalty function to determine “how far” an infeasible individual is from the feasibility constraints of the problem.

We establish two main constraints in our problem. On the one hand, we have a specific number of total charging poles to locate (Ψ) in all different charging stations. Hence, one feasibility function controls that the number of charging poles of an individual (the sum of charging poles of all stations) is specified as the Ψ parameter with an error bound of $\pm 5\%$ (Equation 3):

$$0.95 \cdot \Psi \leq \sum_{\forall cp_i \in \Phi} cp_i \leq 1.05 \cdot \Psi \quad (3)$$

On the other hand, an instance of the problem may have a subset of charging stations with a minimum number of previously installed charging poles as a constraint (this is the set F , as we presented in subsection 3.1). We call these constrained stations *fixed stations*. For this reason, the second feasibility function checks that for each station of the individual, the computed number

of charging poles cp_i is greater than or equal to the number of fixed station charging poles F_i (in the case that there is no fixed charging pole in the station, this number is initialized to 0):

$$\forall i \in \Phi, cp_i \geq F_i \quad (4)$$

These two constraints define the feasibility of the individuals. Any infeasible individual that is generated during the evolutionary algorithm is not desirable as a solution. However, it is important to know the distance to feasibility since just a few mutations or crossover operations can lead to a feasible individual. Therefore, we define the corresponding distance functions for both constraints in order to quantify “how far” an infeasible individual is.

In the case of the total number charging poles, the distance function δ_{cp} is quadratic with respect to the difference of total number charging poles desired and the total number charging poles of the individual.

$$\delta_{cp} = \left(\sum_{\forall cp_i \in \Phi} cp_i - \Psi \right)^2 \quad (5)$$

For the fixed stations, the distance function δ_f is defined as:

$$\delta_f = \sum_{\forall cp_i \in \Phi, cp_i < F_i} (F_i - cp_i)^4 \quad (6)$$

where those stations in which the constraint of Equation 4 is violated, the difference of the minimum number charging poles that should be there (because they are fixed) is raised to the power of 4. This is done to penalize the δ_f distance in a much more severe way than the δ_{cp} distance.

Finally, the function Δ that evaluates the distance to feasibility of an infeasible individual is defined as:

$$\Delta = \delta_{cp} + \delta_f \quad (7)$$

With all of these elements, we can now specify the penalty that is applied to each infeasible individual instead using the fitness function. As we mentioned above, our evolutionary algorithm is implemented using the *deap* library. For this algorithm, there is a mechanism called *Delta penalty* [6] that allows a feasibility function and a distance function to be specified. The feasibility function returns true if the individual is feasible considering the two constraints defined above; otherwise, it returns false. The distance function increases the returned value as the infeasible individual moves away from feasibility.

Therefore, the feasible individuals are evaluated by the fitness function, which considers *utility* and Θ (cost); and the infeasible individuals are evaluated by the delta penalty function, which considers the “distance to feasibility”, as we define below:

$$f(\Phi) = \begin{cases} (utility, \Theta), & \text{if } \Phi \text{ is feasible} \\ (-\Delta, \Delta), & \text{otherwise} \end{cases}$$

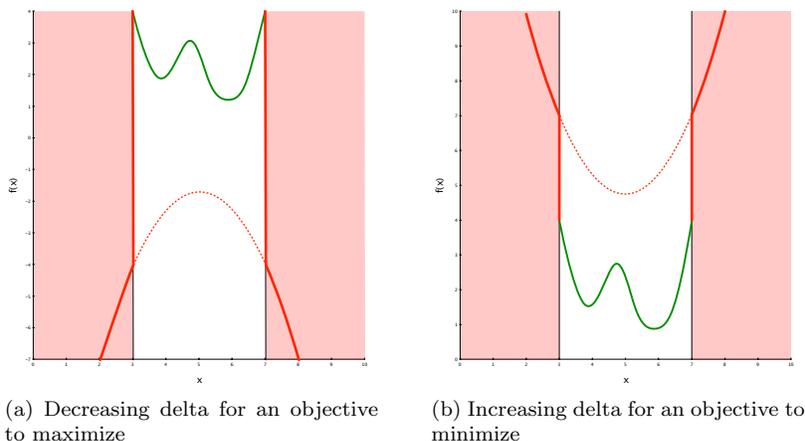


Figure 2: Penalty function examples

When evaluating an individual Φ , the value returned by the algorithm is a tuple with two elements: utility and cost (Θ). However, if Φ is not feasible, the values returned are $(-\Delta, \Delta)$ since the utility must be maximized and the cost minimized. This guarantees that any infeasible individual Φ' will be worse than any feasible individual Φ'' since the utility of Φ' will be a large negative number, while the utility of Φ'' will be a positive number. Similarly, the cost of Φ' will be significantly higher than the cost of Φ'' . An example of this is shown in Figure 2. The left (a) graph shows the possible utility function in green while it is in the feasible central area (top of the graph). The outside areas (light red) represent infeasibility with its corresponding $-\Delta$ function. Similarly, the right (b) graph represents the cost function Θ in green in the feasible area (bottom of the graph), and the Δ function represents the outside zones of feasibility with an increasing value.

5. Case study: City of Lima

In this section, we present a case study of the algorithm proposed in this work for the San Isidro district of the city of Lima (Peru). The city of Lima is very interested in anticipating the future and being prepared for the imminent mass arrival of the EV. For this reason, in this case study, we used the San Isidro district as a benchmark for the algorithm. This is achieved in part thanks to the collaboration of the authorities in providing the necessary data for this case study. The district of San Isidro is one of the 43 districts of the province of Lima. It has an estimated population of 58056 inhabitants, and its total area is $11.1km^2$. It is mainly inhabited by families with a high socioeconomic level.

There is global interest in promoting the EV; for example, the European Commission (EC) has published a *White Paper on Transport* [11] where the EC sets the goals to reach by 2050, betting on the elimination of fossil-fuel based cars.

5.1. Case study example

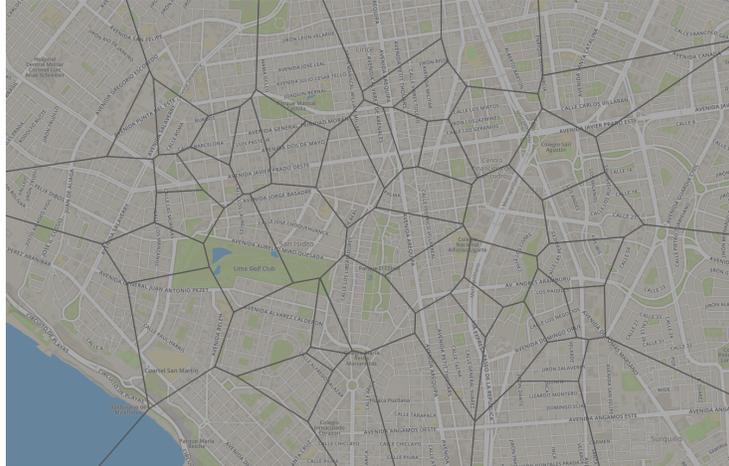


Figure 3: Initial Voronoi diagram based on the PoIs of Lima

For this case study, we have introduced 55 Points of Interest provided by the local district. They have identified places where it would be possible to install an EV charging station. However, let's imagine that the experts of the local government wish to install a significantly smaller number of stations; therefore, the algorithm presented in this work should make a satisfactory and informed proposal. One of the first things that the algorithm performs is to join any PoI that is too close to other PoIs. The minimum distance is selected by the user (e.g., 150 meters) and the PoIs are clustered following this minimum distance. An exception to this situation is when there are existing charging stations, which are treated as a constraint and cannot be removed. In the case of the San Isidro district, there are no previous charging stations in the city, so this constraint does not apply for this case study.

As stated above, to select an area of influence for each Point of Interest, we create a Voronoi diagram using the input PoIs to create regions on the plane of the city. Then, we intersect these regions with a circular area of a diameter chosen by the user (e.g., 600 meters). This gives us the proper area to compute the input variables for the fitness function (i.e., population, traffic, etc.).

This initial situation is shown in Figure 3. As the execution of the genetic algorithm progresses, each selected individual (which represents which stations are active or inactive and with how many charge poles) forces the recalculation of the Voronoi diagram and its intersections with the influence circles (Figure 4). This allows the fitness function to evaluate the quality of that individual (as long as this individual has previously passed the feasibility check). Note that an individual will not be considered feasible when it has more charge poles than the number accepted by the user or when it tries to reduce the number of charge poles of a previously installed station.



Figure 4: Example of the Voronoi diagram and influence circles of each PoI in Lima

Finally, the genetic algorithm returns a set of possible solutions with a good balance between utility and cost. Figure 5 shows a possible solution for this case study. In this solution, it can be observed that the number of points selected as candidates for installing charging stations is significantly lower than the initial number of Points of Interest proposed as inputs. For this reason, the set of Voronoi polygons shown in the solution is different from the initial one (Figure 3) because they are adapted and recalculated for each solution evaluated.

One approach for seeing how the solution has been adapted to the characteristics of the city is to check how the charging stations are distributed in relation to the electrical transformer substations. The algorithm has tried to minimize the cost of deploying charging stations that are too far from these substations. This is why Figure 6 shows how the distribution of the stations resembles the coverage of the transformer substations, which is shown as a heatmap. Similar checks could be done by comparing the solution with the distribution of population, traffic, or social activity.

5.2. Experimental Results

In this subsection, we present some of the experimental results of our proposal when applied to the city of Lima. As mentioned above, our NSGA-II evolutionary algorithm is implemented using the *deap*² library of *Python*. The settings that we use in these experiments are summarized in Table 1.

For the main parameters of the genetic algorithm, we considered different initial populations ranging from 100 to 5000 individuals, an evolution through 300 generations, and the probabilities for the crossover and mutation operations of 0.5 and 0.05, respectively. The values $(\mu + \lambda)$ of the NSGA-II algorithm are

²<https://github.com/DEAP/deap>

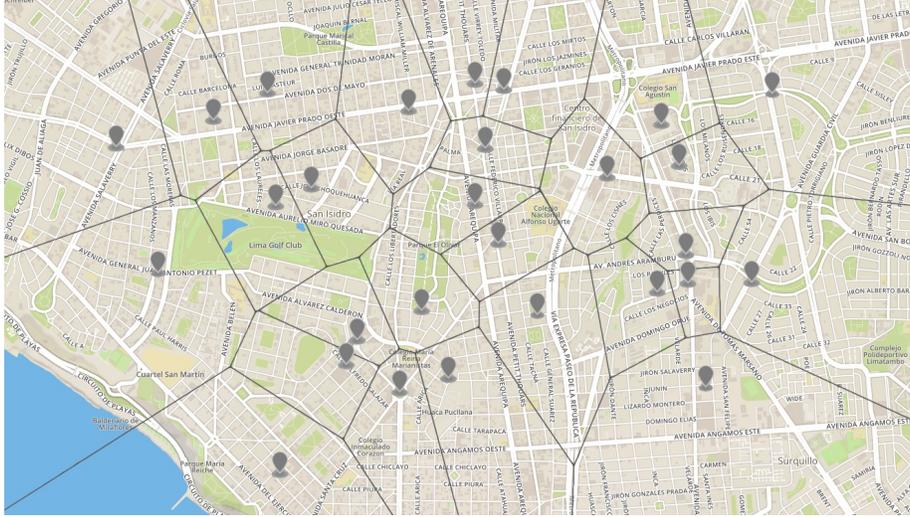


Figure 5: EV charging stations location solution for the Lima case study example

defined depending on the population. Therefore, the number of individuals to select at each generation is $\mu = \text{population}$, and the number of descendants to produce at each generation is $\lambda = \frac{\text{population}}{2}$.

The main parameters of these experiments for the San Isidro district were the following. We increased the input set of PoIs from 55 to 548 by including additional PoIs to the original 55 PoIs provided by the authorities, which we show in the above case study. This implies that the problem has a larger search space, which increases the probabilities of a genetic algorithm finding a better solution by expanding that search space. With this greater number of PoIs, the minimum distance among PoIs is established as 50 meters, and, hence, all PoIs at a closer distance are clustered. This clustering process reduces the original 548 PoIs to 379, with a minimum distance among them of 50 meters. Taking into account the shape, the characteristics, and the population of the San Isidro district, the number of charging poles to locate in this problem is 50, with a maximum of 3 charging poles per station. We considered this setting to be the most appropriate for the district regarding the needs for an EV charging infrastructure.

We have already provided above the details and formulas that we used to calculate the values of fitness, which is based on the following parameters. The weights used in the utility function were: $\omega_P = 0.4$, $\omega_T = 0.35$, and $\omega_A = 0.25$ since we consider population to be the most important parameter, followed by traffic. The costs that affected the fitness values were: the cost of placing each charging station (10000); the cost per charging pole (40000) (note that the cost of placing the station is, in fact, the cost of placing the first charging pole, which is more expensive than the additional poles); and the cost per meter to connect the nearest transformer substation to the charging station



Figure 6: Heatmap of energy transformer substations in Lima compared to the solution of the case study.

(150). The limit to connect to a transformer substation was 100 meters, which means that any longer distance would double this cost. Finally, the influence radius of the charging stations was 100 meters since the minimum distance among PoIs was 50 and the number of PoIs with respect to the previous case study was considerably higher. It is important to note that these values were obtained through consultations with companies that specialize in the installation of charging points for EVs.

We consider that while the genetic algorithm is configurable, the rest of the parameters are a setting of the problem that is not directly comparable with a different one. Hence, our experiments focused on analyzing the quality of the solutions obtained by the variation of the genetic algorithm configuration with the same problem setting.

Figure 7 shows the evolution of the average utility of individuals through generations of the problem setting presented in Table 1 with different initial populations for the genetic algorithm. Each line represents the average of 10 different executions with the corresponding initial population, and it shows the average utility of the individuals in the population at each generation. Globally, the utility remains in relatively small values since it represents the percentage of area covered by the charging stations. In order to have full coverage of the whole area, all PoIs must have a charging station and the union of the influence areas of all PoIs should totally cover the areas in which there is population, traffic, and social activity. Therefore, it is unlikely that such a solution exists, but also, in the case that concerns us with only 50 poles to locate, it is impossible to cover the whole area. That is why we simply seek maximum coverage with the provided resources.

Table 1: Configuration parameters for the Lima case study

Parameter	Value
population	100-5000
generations	300
crossover prob.	0.5
mutation prob.	0.05
number of PoIs	548
distance PoIs	50
total poles to locate	50
max poles per station	3
ω_P	0.4
ω_T	0.35
ω_A	0.25
cost station	10000
cost pole	40000
cost distance energy	150
energy radius	100
influence radius	100

All of the utility values in Figure 7 are positive since there is no infeasible individual generated in the genetic algorithm. On the one hand, we used a controlled method to generate the initial population to avoid creating infeasible individuals. On the other hand, the possible infeasible individuals created through the generations of the evolutionary algorithm are not selected for next generations since there are enough feasible individuals that are much better than those infeasible individuals that have a high penalty of $-\Delta$ for utility. Also note that, as we explained in Section 4.3, the only infeasible individuals that can be produced are those with a number of charging poles that differs more than 5% of the specified value of 50 in these experiments, since we do not consider any previous existing charging station in the San Isidro district.

The lines of the graph in Figure 7 clearly follow an increasing curve through generations with small oscillations. The experiments with an initial population of 100 individuals have an average utility for their individuals that is significantly lower than the other experiments. Additionally, the average utility of the experiments with 100 individuals of initial population hardly improves from the 100 generations and it stabilizes when arriving at 150 generations. This occurs because such a small population prevents the algorithm from evolving beyond these generations; this phenomenon is not so evident in other experiments with a larger population.

For the rest of the initial populations, the evolution is quite similar and the significant differences of average utility among them begin at 150 to 200 generations. The experiments with 250 individuals in the population stabilize

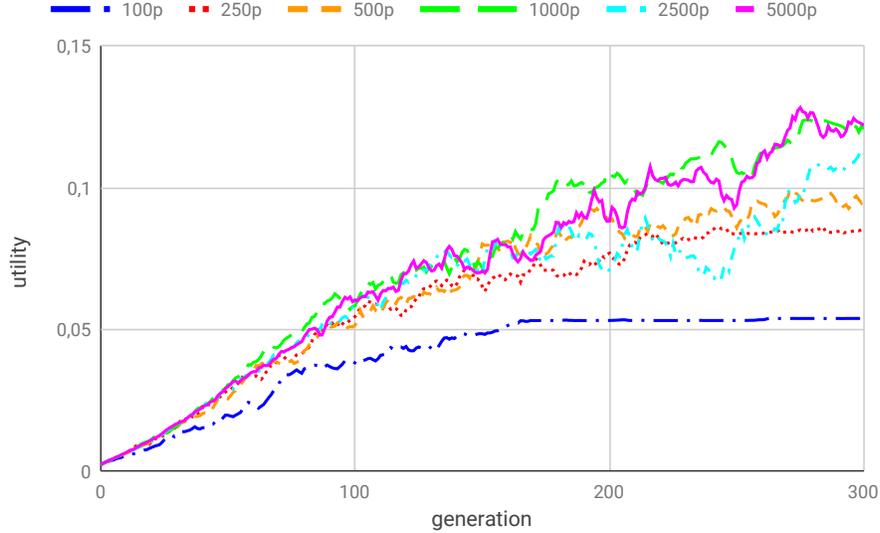


Figure 7: Evolution of average utility starting with different populations

at around generation 250, while the experiments with a higher population still vary. At generation 300, the average utilities of the experiments are ordered from the highest to the lowest initial population, with the experiment with a population of 5000 being the one with the highest utility. However, there is an exception in the experiments with a population of 1000 individuals since they have the highest global average utility in many generations and also a higher average utility than the experiments of 2500 individuals even at generation 300. It is worth noting that in this graph, we are analyzing the average utility of all of the individuals in the population, and, for this reason, there may be many individuals that decrease or increase this value. Therefore, in order to be able to observe how the number of individuals in the initial population influences the utility when generating the best individuals, we represent the graph of the maximum utility values for these experiments.

The graph in Figure 8 shows the maximum utility among all of the individuals at each generation for several numbers of individuals in the initial population of the NSGA-II algorithm. As in Figure 7, these results are the average of 10 executions for each initial population. This graph shows how maximum utility grows significantly in the first 100 generations in all cases. As we observed above, the case with 100 individuals stabilizes at around 125 generations and it is unable to improve the maximum utility for any individual. Therefore, for this case, 150 generations would be enough. A similar behaviour is observed for the cases of 250 and 500 individuals of initial population (approximately at generation 175 and 200, respectively) with just some small variation in the case of 500 individuals in the final 50 generations. For the cases of 1000, 2500, and 5000

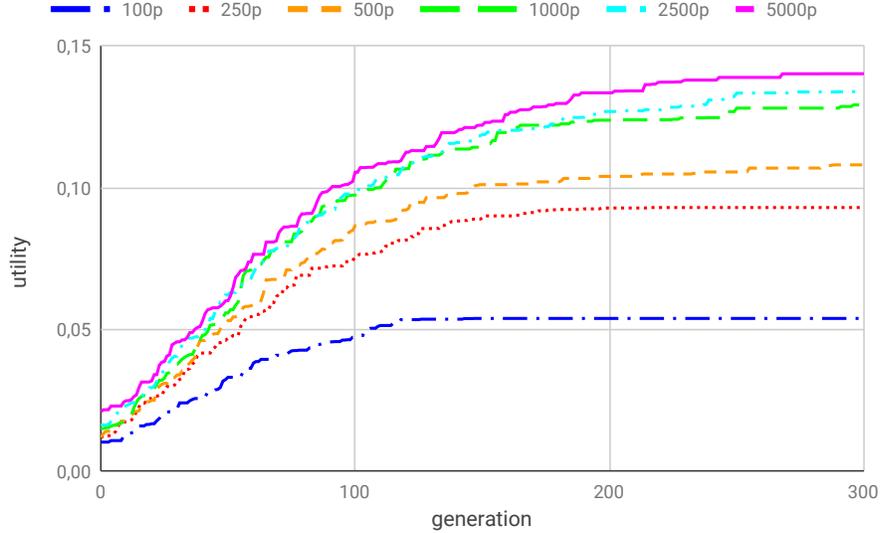


Figure 8: Evolution of maximum utility starting with different populations

individuals, the maximum utilities reached have a very close value and also have a very similar evolution so that the maximum utility almost does not increase in the last 50 generations. Therefore, we can consider that 300 generations are enough to obtain a solution for these problem settings.

The maximum utility values shown in Figure 8 are ordered with respect to the initial population, with the 5000-individual experiments being the ones with the highest maximum utility. Finally, it should be noted that the difference between the maximum utility values of the experiments with less than 1000 individuals of initial population is significant with respect to those of 1000 individuals or more. However, there is no significant difference in maximum utility between the values of the experiments with 1000, 2500, and 5000 individuals. Therefore, given the complexity of the execution of 5000 individuals with respect to 1000, and the small difference in the quality of the final solution, it is up to the user to determine how many individuals between these ranges a configuration should have in order to reach a satisfactory solution.

A similar graph that plots the cost Θ could be done since cost behaves similarly to utility, but with different scales and orders of magnitude. Note again that the multi-objective genetic algorithm works perfectly in the sense that we can optimize objectives that represent totally different values because we do not have to put them together in the same formula leveraging them.

Figure 9 compares the evolution of utility and cost in order to show how the NSGA-II algorithm performs with the multi-objective function that maximizes utility and minimizes cost Θ . The lines represent the average utility (left axis) and the cost (right axis), respectively, of the averaged results of the 10 repeti-



Figure 9: Evolution of average utility and cost

tions for the experiment with 1000 individuals of initial population. It can be observed that there is great improvement in both utility and cost in the first 100 generations, and then the cost is stabilized while the utility keeps increasing. This indicates that the algorithm is able to find a stable cost using all resources to locate the 50 charging poles, and then it keeps generating individuals with a similar cost but changing the distribution of the stations to increase the coverage, thereby increasing the utility.

Since the solutions returned by the algorithm are at the Pareto frontier, i.e., the objectives are optimized independently and there are none that are Pareto dominated, there are several alternatives to be chosen by the user. In the tests carried out, the users must choose between several solutions in which if we consider a solution Φ and an alternative Φ' , the utility of Φ is greater than that of Φ' , while the cost Θ of Φ is in turn greater than that of Φ' . Hence, it is up to the users, who should be experts in the field of electric mobility, to choose the alternative that best suits their needs. Although it must be noted that the differences between the solutions are not very significant, what is really significant is the reduction from the possible initial combination to the set of possible solutions found by the genetic algorithm. This provides decisive guidance during the decision-making process.

The obtained solutions tend to have one charging pole per station to cover the demand, except for some stations located in central areas with more population, traffic, and social network activity that have more than one charging pole. Hence, the solutions have around 45-50 stations out of the 548 initial POIs

clustered to 379, and 48-52 charging poles of the 50 ($\pm 5\%$) to be located initially. Therefore, we can consider these results satisfactory since the algorithm behaves as expected.

Figure 10 shows a solution obtained with an initial population of 5000 in one of the 10 repetitions of the experiment. In this solution, the charging stations are distributed quite equally in the San Isidro district in an attempt to cover the area, but they are more concentrated in the center of this district since there is more population and social activity there. It is important to note that this solution differs from the one in Figure 5 regarding the size of the Voronoi polygons since the PoIs are closer and their influence areas are also smaller. Finally, it must also be noted that the solutions obtained in different repetitions of the same setting have very similar utility and cost values, and these solutions have just slight differences in their shape, with some of the stations varying, with the ones represented in Figure 10.



Figure 10: Solution for the location of EV charging stations in Lima

6. Conclusion

In this paper, we have presented a multi-objective genetic algorithm to solve the problem of selecting the most suitable locations for installing electric vehicle (EV) charging stations in a smart city. The algorithm takes into account extracted information extracted about the city. This includes population distribution, traffic intensity, activity in social networks and the location of energy transformer substations as well as the cost of installing the stations and the number of charging poles for each station to cover the demand.

We have improved previous works by introducing a multi-objective evolutionary algorithm that optimizes both utility and cost, generating solutions

along the Pareto frontier. This kind of algorithm allows us to find a solution that finds a balance between the greatest benefit of installing the proposed stations in the solution and the lowest possible cost from among the solutions of this type. Thus, the decision-maker in the deployment of an EV charging infrastructure in a city can use this information to maximize the benefit of the deployment without excessively raising the cost or wasting resources. We have also included the energy transformer substations of the city in the algorithm in order to determine the best locations for the new EV charging stations, taking into account the existing power infrastructure in order to rationalize the connections and impact on the city. This feature also provides a better estimate of the cost of deployment since EV charging stations have to be connected to transformer substations. Thus, the closer the charging stations are located to the transformer substations, the lower the cost of supplying the energy to the chosen points. Another of the improvements proposed by this work is the use of social activity information to not only track where users live in a city, but also where they move throughout the day. This information is crucial in determining where there is a greater likelihood of having users who need to charge their EVs throughout the day. In addition, since we only use the intensity of this social activity in order not to overload the evolutionary algorithm with the enormous amount of geo-localized information existing nowadays, the solution has been optimized using buckets where the information has been discretized. This has improved the performance of the evolutionary algorithm. Finally, the system was analyzed through a case study in the city of Lima (Peru), and it was empirically evaluated.

As future work, the energy capacity transformer substations can be added as a constraint of the problem to perfectly fit the charging stations to be located by taking into account the available capacity, thus reducing the impact on the current energy infrastructure. In addition, a search will be made for the best algorithm configuration parameters in order to maximize the quality of the results. Furthermore, the users of the system (i.e., administrations, governments, or city council managers) might want to prioritize certain PoIs or city zones that are of special importance and that are not totally reflected in the available data. This would make our system more adaptive to the requirements of the city managers in order to strategically deploy an adequate charging infrastructure. Finally, it would be useful to improve our system with a technique that plans a phased installation over several years of the charging infrastructure. This way, the introduction of EV charging stations could be done incrementally to cover the demand as it grows, while rationalizing the deployment costs over time.

Acknowledgments

This work was partially supported by MINECO/FEDER RTI2018-095390-B-C31 project of the Spanish government. Jaume Jordán and Vicent Botti are funded by UPV PAID-06-18 project. Jaume Jordán is also funded by grant APOSTD/2018/010 of Generalitat Valenciana - Fondo Social Europeo.

References

- [1] Alternative Fueling Station Locator. <http://www.afdc.energy.gov/locator/stations/>, 2019. Accessed on 17-01-2019.
- [2] Community Energy Association. Zap Map Toolkit. <https://www.zap-map.com/live/>, 2019. Accessed on 17-01-2019.
- [3] Charge Hub. <https://chargehub.com/en/charging-stations-map.html>, 2019. Accessed on 17-01-2019.
- [4] Chargepoint. https://eu.chargepoint.com/charge_point, 2019. Accessed on 17-01-2019.
- [5] Charging Station Map Worldwide. <https://www.plugshare.com/>, 2019. Accessed on 17-01-2019.
- [6] Carlos A Coello Coello. Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. *Computer Methods in Applied Mechanics and Engineering*, 191(11):1245 – 1287, 2002.
- [7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, April 2002.
- [8] Jing Dong, Changzheng Liu, and Zhenhong Lin. Charging infrastructure planning for promoting battery electric vehicles: An activity-based approach using multiday travel data. *Transportation Research Part C: Emerging Technologies*, 38:44–55, 2014.
- [9] EV Go. <https://www.evgo.com/charging-locations/>, 2019. Accessed on 17-01-2019.
- [10] EV Infrastructure Corridor Development Toolkit. <http://altfueltoolkit.org/ev-infrastructure-corridor-development-toolkit/>, 2019. Accessed on 17-01-2019.
- [11] Directorate General for Mobility and Transport. Roadmap to a single European Transport Area - Towards a competitive and resource-efficient transport system. https://ec.europa.eu/transport/sites/transport/files/themes/strategies/doc/2011_white_paper/white-paper-illustrated-brochure_en.pdf, 2011. Accessed on 30-01-2019.
- [12] Thomas Franke, Madlen Günther, Maria Trantow, Josef F Krems, Roman Vilimek, and Andreas Keinath. Examining user-range interaction in battery electric vehicles—a field study approach. *Advances in human aspects of transportation part II*, pages 334–344, 2014.

- [13] Thomas Franke and Josef F Krems. Understanding charging behaviour of electric vehicle users. *Transportation Research Part F: Traffic Psychology and Behaviour*, 21:75–89, 2013.
- [14] Lingwen Gan, Ufuk Topcu, and Steven H Low. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems*, 28(2):940–951, 2013.
- [15] Google Maps Application. <https://maps.google.com>, 2019. Accessed on 17-01-2019.
- [16] Jia He, Hai Yang, Tie-Qiao Tang, and Hai-Jun Huang. An optimal charging station location model with the consideration of electric vehicle’s driving range. *Transportation Research Part C: Emerging Technologies*, 86:641 – 654, 2018.
- [17] Jaume Jordán, Javier Palanca, Elena del Val, Vicente Julian, and Vicente Botti. Using genetic algorithms to optimize the location of electric vehicle charging stations. In *International Joint Conference SOCO’18-CISIS’18-ICEUTE’18*, pages 11–20. Springer International Publishing, 2019.
- [18] Jaume Jordán, Javier Palanca, Elena Del Val, Vicente Julian, and Vicente Botti. A multi-agent system for the dynamic emplacement of electric vehicle charging stations. *Applied Sciences*, 8(2), 2018.
- [19] Raja Jurdak, Kun Zhao, Jiajun Liu, Maurice AbouJaoude, Mark Cameron, and David Newth. Understanding human mobility from twitter. *PloS one*, 10(7):e0131469, 2015.
- [20] Gillian Lacey, Ghanim Putrus, Edward Bentley, David Johnston, Sara Walker, and Tianxiang Jiang. A modelling tool to investigate the effect of electric vehicle charging on low voltage networks. In *Electric Vehicle Symposium and Exhibition (EVS27), 2013 World*, pages 1–7. IEEE, 2013.
- [21] Zhenping Li and Xiaojie Cui. Research on Location Problem of Electric Vehicle Charging Station. *Journal of Applied Science and Engineering Innovation*, 2(12):495–498, 2015.
- [22] Tan Ma and Osama A Mohammed. Optimal charging of plug-in electric vehicles for a car-park infrastructure. *IEEE Transactions on Industry Applications*, 50(4):2323–2330, 2014.
- [23] Zachary A Needell, James McNerney, Michael T Chang, and Jessika E Trancik. Potential for widespread electrification of personal vehicle travel in the United States. *Nature Energy*, 1:16112, 2016.
- [24] Yu Marco Nie and Mehrnaz Ghamami. A corridor-centric approach to planning electric vehicle charging infrastructure. *Transportation Research Part B: Methodological*, 57:172–190, 2013.

- [25] Gustavo Niemeyer. Geohash. <http://geohash.org/site/tips.html>, 2008. Accessed on 23-01-2019.
- [26] Open Charge Map. <https://map.openchargemap.io/>, 2019. Accessed on 17-01-2019.
- [27] Planning for Electric Vehicle Charging Infrastructure: A Toolkit. <http://pluginbc.ca/resource/planning-electric-vehicle-charging-infrastructure-toolkit/>, 2019. Accessed on 17-01-2019.
- [28] Aviral Shukla, Joseph Pekny, and Venkat Venkatasubramanian. An optimization framework for cost effective design of refueling station infrastructure for alternative fuel vehicles. *Computers & Chemical Engineering*, 35(8):1431–1438, 2011.
- [29] Timothy Sweda and Diego Klabjan. An agent-based decision support system for electric vehicle charging infrastructure deployment. In *Vehicle Power and Propulsion Conference (VPPC), 2011 IEEE*, pages 1–5. IEEE, 2011.
- [30] Wei Tu, Qingquan Li, Zhixiang Fang, Shih-lung Shaw, Baoding Zhou, and Xiaomeng Chang. Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach. *Transportation Research Part C: Emerging Technologies*, 65:172–189, 2016.
- [31] Sebastian Wagner, Markus Götzinger, and Dirk Neumann. Optimal location of charging stations in smart cities: A points of interest based approach. In *Thirty Fourth International Conference on Information Systems*, Milan, Italy, 2013. ICIS.
- [32] Eric Wood, Jeremy S Neubauer, and Evan Burton. Measuring the benefits of public chargers and improving infrastructure deployments using advanced simulation tools. Technical report, SAE Technical Paper, 2015.