This is the accepted manuscript of: Batic, D., Stankovic, V., & Stankovic, L. (in press). ChargeDEM: geodemographic aware EV charging infrastructure placement for enhanced site selection using graph neural networks. In *12th International Conference on Energy Efficiency in Domestic Appliances and Lighting (EEDAL'24)*

ChargeDEM: Geodemographic Aware EV Charging Infrastructure Placement for Enhanced Site Selection using Graph Neural Networks

Djordje Batic, Vladimir Stankovic, Lina Stankovic

Department of Electronic and Electrical Engineering, Faculty of Engineering, University of Strathclyde, Glasgow, UK

Abstract

Electric vehicles (EVs) have become a key factor in the shift towards sustainable transportation. Yet, the rapid growth in EV adoption has outpaced the development of adequate EV charging infrastructure, leading to a critical adoption bottleneck. While recent studies have increasingly focused on optimizing charging station placement through mathematical modelling and decision-making strategies, they often overlook the intricate spatial dynamics among charging demand nodes. Moreover, the impact of new charging stations on the utilization of existing charging infrastructure is rarely accounted for in the site selection process. Additionally, socio-demographic factors are frequently neglected, potentially marginalizing underserved communities. To address these critical gaps, we propose a novel, geodemographic aware approach for EV charging site selection. This method leverages graph neural networks (GNNs) to identify optimal locations for charging stations while maximizing the efficiency of the installed charging infrastructure. We apply our approach to a case study of the Glasgow City area, Scotland, UK, demonstrating the potential to effectively guide infrastructure planning. The methodology not only significantly reduces installation costs but also boosts the utilization of urban charging facilities. By considering socio-demographic, spatial, and post-installation factors, this approach offers a holistic solution for the fair and efficient growth of EV charging infrastructure.

1 Introduction

In the wake of escalating environmental concerns and the pressing need to mitigate climate change, the transportation sector has been identified as one of the cornerstones of a successful transition to net zero [1]. The longstanding reliance on vehicles that utilize common fossil fuels, while integral to societal advancement, has significantly contributed to greenhouse gas emissions and environmental degradation [1]. In recent years, alternative fuel vehicles, such as Electric Vehicles (EVs), have emerged as a transformative technology in this context, offering a cleaner, more sustainable mode of transportation. Global penetration of alternative fuel vehicles has experienced rapid growth in recent years [2]. In the UK alone, at the end of 2023, there were over 735,000 registered Battery Electric Vehicles (BEVs) and over 480,000 Plug-In Hybrid Electric Vehicles (PHEVs), a number expected to sky rocket in the coming years [3]. However, despite an optimistic market expansion forecasts, rapid growth in EV ownership is juxtaposed against a lagging EV charging infrastructure, creating a critical adoption bottleneck [4]. Development of accessible, convenient and reliable charging infrastructure is not keeping pace with the surge of EVs on the road. At the same time, the optimal placement of charging infrastructure is a complex problem, influenced by a myriad of factors including regional uptake of EVs, urban planning constraints, electrical grid capacity, accessibility, and consumer behavior patterns. Similarly, utilization of deployed charging stations is dependent on various factors, including accessibility, position within the road network, proximity to areas of public interest, area coverage, reachability, safety, area traffic flow, and connectivity to the public transport network [5]. Lastly, ensuring fair access to charging facilities remains a key concern, guaranteeing that all members of the public, regardless of geographical location or sociodemographic group, can partake in the shift towards more sustainable transportation.

In the realm of EV charging infrastructure planning, various modeling methods, optimization techniques, and decision-making approaches have been introduced [6]–[9]. These methodologies can generally be classified into two categories: flow-based and node-based models. Early research tackled the challenge of predicting charging demand using a flow-based perspective. which assumes that traffic volume near a prospective site is the primary determinant of charging demand, and represents charging demand as a series of origin-destination trips [10]. However, this approach has limitations, particularly in urban settings, as it fails to account for other significant elements influencing urban vitality [11]. Conversely, node-based approaches favor the view that charging demand is dependent on demand nodes within the urban network. This perspective suggests that EV drivers prefer charging locations that are conveniently located, such as near their homes, workplaces, or other key points of interest like public service facilities and recreational areas. While node-based approaches provide valuable insights into urban charging requirements, they often neglect the broader socio-demographic and spatial dynamics that play a crucial role in determining the feasibility and effectiveness of EV charging infrastructure deployment.

To overcome the challenges in infrastructure placement, it is essential to consider a multidimensional set of data points in order to accurately model charging demand. This approach involves integrating demographic data, such as human development indices, information on points of interest (POIs), charging infrastructure information, traffic flow, and geospatial topology. By adopting this multi-dimensional methodology, we facilitate a more nuanced understanding of the potential utilization and effectiveness of EV charging infrastructure. Specifically, utilizing geodemographic data allows for the consideration of socio-economic factors that influence EV adoption rates and charging needs, which ultimately can help strategic placement of chargers for underserved communities to ensure fair access and encourage EV adoption. Placement decisions are often made without understanding the spatial dynamics of the urban network, an often-overlooked aspect of existing node-based approaches. In this paper, we propose a geodemographic, spatially-aware approach for EV charging station placement using graph neural networks (GNNs), which ensures that the deployment and expansion of EV charging infrastructure is both efficient and fair, addressing the critical bottleneck in EV adoption, and contributing to the larger goal of transitioning to more sustainable transportation. We conduct thorough experiments using real-world data from the selected urban site of the city area of Glasgow, Scotland, UK. As a major urban center and the largest city in Scotland - a region known for its high renewable energy production and lower than average national carbon footprint [12] - Glasgow presents an ideal environment for advancing EV usage with a reduced carbon footprint.

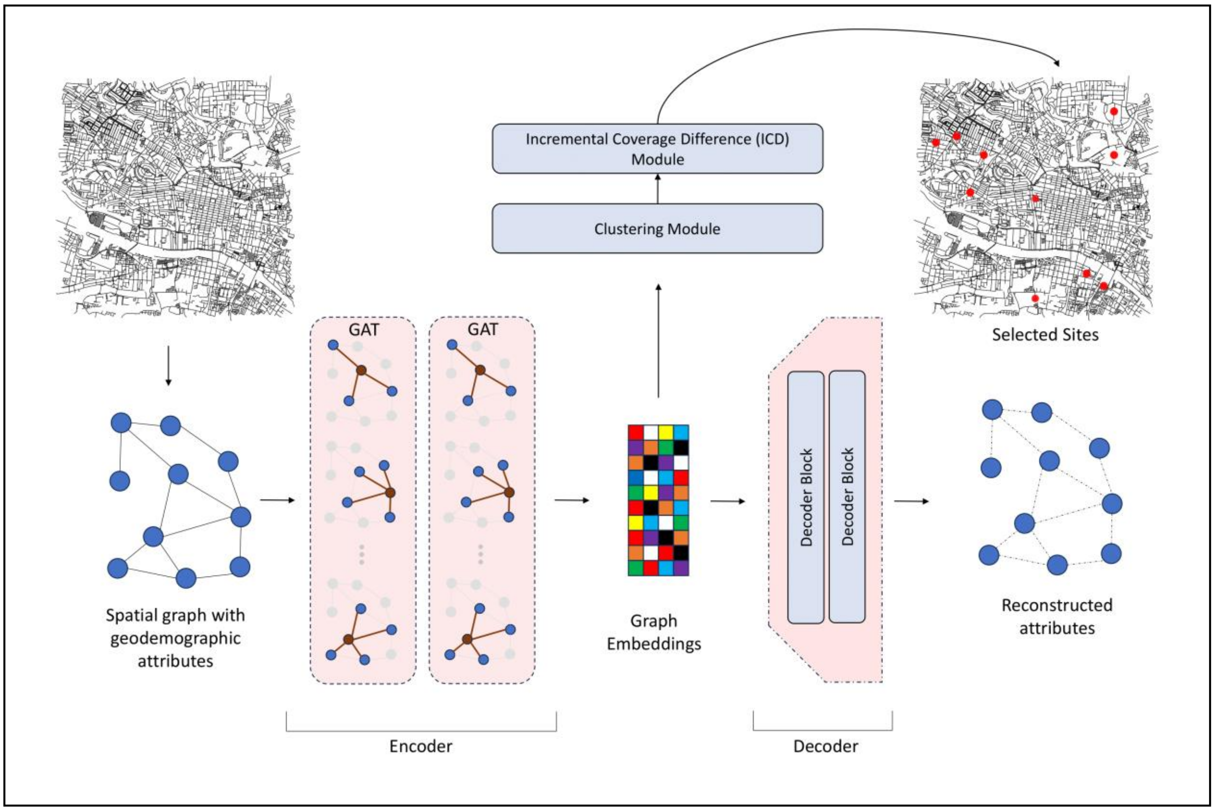
The rest of this paper is organized as follows: Section 2 introduces the related work of EV charging site selection. Section 3 illustrates the proposed graph representation learning mechanism design framework. Section 4 presents the experimental study. Section 5 draws conclusions and suggests future research.

2 Background

Traditionally, the problem of EV charging station site selection has been approached as a decision-making process that aims to identify an optimal site based on a set of predefined criteria. This approach hinges on carefully defined criteria by experts or sophisticated mathematical modelling. Approaches include the use of mixed integer linear programming (MILP) [6], multi-objective programming (MOP) [7], and algorithmic approaches such as genetic algorithms [8]. However, as selection criteria contains multiple (subjective) factors that often have conflicting properties, approaches that take this important aspect into account have been proposed. Among them, multi criteria decision making (MCDM) are one of the most popular site selection methods. MCDM has been used in various site selection domains, such as offshore wind farms [13], emergency hospital placement [14], and landfill sites [15]. MCDM follows a hierarchical process that evaluates and prioritizes different selection objectives across candidate sites to determine the most suitable locations. A key aspect of MCDM is the assignment of weights to various criteria, which can be subjective and influenced by the decisionmaker's preferences or biases. Moreover, MCDM typically relies on static objectives and lacks the capability to capture complex relationships within data. This limitation is particularly pertinent in contexts like EV charging station placement, where modeling of the spatial dynamics between evolving and interdependent charging demand hotspots is crucial for effective site selection.

To enhance the understanding of spatial relationships between urban demand nodes and their influence on site placement, recent research has investigated the use of GNNs in the urban mobility domain. [16] utilized GNNs to predict the attractiveness of different store sites within neighborhoods based on public transport hotspots, [17] propose a spatiotemporal graph convolutional network for traffic forecasting, combining graph convolutions and gated temporal convolutions to extract the relevant spatial features and capture key temporal dynamics. However, despite these successes in the urban mobility field, the application of GNNs for EV charging infrastructure placement has been relatively unexplored. The first known approach using GNNs for this purpose uses GNNs to select candidate sites of charging stations, where the sites are evaluated using traffic flow and construction cost information [9]. This method, however, divides the data into equally sized grid cells, imposing artificial boundaries that might not correspond with natural charging demand nodes. It also assumed uniformity within each cell, which limits the ability to utilize socio-demographic characteristics of demand nodes. Additionally, this approach did not consider how new stations might affect existing charging demands within each node, a critical aspect in determining the most effective installation sites.

Figure 1 – Proposed ChargeDEM EV charging site selection approach.



3. Methodology

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represent a graph with $|\mathcal{V}|$ number of nodes and $|\mathcal{E}|$ number of edges, where $\mathcal{X} \in R^{|\mathbb{V}| \times F}$ denotes a set of features of all nodes, where F is the feature dimension. In particular, tackling the EV charging station placement problem, we consider a charging demand node $v \in \mathcal{V}$ to be the area with radius r centered around an existing EV charger site, while edges $\mathcal{E}$ represent undirected connections between nearby nodes going from node $u \in \mathcal{V}$ to node $v \in \mathcal{V}$ such that $(u, v) \in \mathcal{E} \leftrightarrow (v, u) \in \mathcal{E}$. That is, a pair of nodes in the graph are connected if their physical distance is at most r. Let $A \in \{0,1\}^{|\mathcal{V}| \times |\mathcal{V}|}$ represent the graph adjacency matrix with entries $a_{u,v} = 1$ if there is an edge between nodes u and v, and $a_{u,v} = 0$,otherwise, all pairs of nodes $u, v \in \mathcal{V}$, and $a_{u,u} = 0$.

In this work, we argue that to successfully extrapolate this graph-based approach in the realm of EV charging station placement, it is crucial to note that each node in $\mathcal{G}$ is not merely a point in space but encapsulates a comprehensive data-driven representation of its surrounding environment. This representation includes, but is not limited to, the sociodemographic makeup, density and type of POIs, traffic patterns, and existing parking infrastructure within the confines of r. The edges $\mathcal{E}$, on the other hand, embody the physical and functional connectivity between these nodes, influencing the flow and demand of EV charging within the network. The integration of this diverse set of data within the graph structure allows for a nuanced understanding of the EV charging demand dynamics at each node. To achieve this, we place the underlying problem of capturing structural information and encoding of relational context of urban charging demand in the framework of graph representation learning through graph autoencoders. This approach will enable effective capture and utilization of complex relationships inherent in the graph structure of urban charging demand that include diverse influencing factors such as spatial characteristics, existing demand, urban vitality, demographic makeup, traffic flow, reachability, safety. This approach also benefits from accounting for relationships between neighboring demand nodes, an often-overlooked aspect of charging station site selection.

An overview of the proposed sociodemographic, spatially aware GNN based framework for EV charging infrastructure placement is shown in Figure 1. The approach consists of three parts: 1) graph representation learning for capturing complex relationships inherent in the graph structure of urban charging demand 2) clustering of node representations and subsequent identification of low potential, medium potential, and high potential installation areas 3) utility-informed site selection based on area potential, taking into account installation requirements, such as the number of chargers to be installed, their installation utility, and how they affect the overall charging potential of surrounding areas. In this section, we describe each component in detail.

3.1 Urban charging graph representation learning

Use of autoencoders for learning graph representation without supervision based on graph neural networks (GNNs) has been proposed in [18]. An autoencoder usually contains an encoder, latent representations, and a decoder. The encoder aims to map the input data to latent representations, and leverages on latent representations to reconstruct the input under supervision of a reconstruction criterion. Given f_e as the graph encoder, f_d as the graph decoder, and H' representing the encoder latent representations, the aim of graph autoencoders is to learn the mapping:

$$H' = f_e(A, X), G' = f_d(A, H') \tag{1}$$

where G' represents the reconstructed features or structure of the graph by the decoder.

The use of completely symmetric graph convolutional autoencoders which utilize both the structure of the graph and node attributes through the whole encoding-decoding process has been proposed in [19]. In such approaches, the issue of instability of common graph convolutional layers is tackled by the addition of Laplacian sharpening layers to counterbalance the common smoothing effects. However, even though this approach fixes common graph representation learning problems, it does not allow for dynamic weighting of node importance weights. In this setting, GNNs use a set of node features to learn node representation. Typically, GNNs perform a node-level feature aggregation within a neighborhood, which involves iterative learning of node representation through aggregation of representations of its neighbors to create H' latent representation. However, this approach usually assumes equal importance of all neighbors and assigns the same aggregation weights based on the degree distance.

To tackle this problem, we leverage on and adapt the common approach of employing the self-attention mechanism in the encoding step using Graph Attention Network (GAT) [20]. Unlike typical GNNs, GATs assign dynamic weights to neighboring nodes based on relative importance of nodes in the neighborhood through masked attention mechanism. Given the input as a set of node features $H = \{H_1, H_2, \dots H_{|\mathcal{V}|}\}, H_i \in R^F$, a graph attention layer computes output $H' = \{H'_1, H'_2, \dots H'_{|\mathcal{V}|}\}, H'_i \in R^{F'}$ with a different cardinality F'. To compute the weights, masked self-attention mechanism is performed, where attention coefficient α_{ij} that indicate the importance of features in node j to node i are computed only for nodes $j \in \mathcal{N}_i$ where $\mathcal{N}_i$ is some neighborhood of node i in the graph:

$$\alpha_{ij} = \frac{\exp\left(a\left(WH_i, WH_j\right)\right)}{\sum_{k \in \mathcal{N}_i} \exp\left(a(WH_i, WH_k)\right)} \tag{2}$$

where a represents attention function, and W a weight matrix, where $W \in R^{F' \times F}$. Additionally, as an urban graph tends to have large number of neighbors, in order to mitigate "neighbor explosion" problem, the data sampling procedure is performed by obtaining a set of subgraphs by sampling the original training graph and then building the graph autoencoder based on the subgraph [21].

The ability to assign dynamic node weights is a crucial step in modelling the GNN approach to the charging station placement setting where model allows for the implicit assignment of different importance weights to nodes of a same neighborhood, enabling representation learning of graphs that encompass dense information sources such as traffic patterns, street topology, demographic factors, charging demand zones, urban mobility etc. ChargeDEM EV charging site selection approach is illustrated in Fig. 1. Contrary to the aforementioned traditional graph representation learning approach, which makes use of common graph convolutional network (GCN) spectral convolutional layers, ChargeDEM consists of a GAT-based encoder and a simplified, linear decoder layer for reconstruction of graph representation. This formulation enables dynamic weighting of node features and strikes a balance between model learning capacity through self-attention and smoothing minimization.

3.2 Clustering

After acquiring graph embeddings that reflect the underlying structure of urban charging demand, informed by factors inherent in the node features, we harness this data to pinpoint areas with high charging potential. Our dataset lacks explicit information, such as historical charger utilization records, making it impossible to frame this as a supervised problem. Instead, we adopt an unsupervised clustering approach, similar to one used in previous research [9], to identify charging nodes with significant potential. To determine the optimal number of clusters, we use the K-means [22] clustering combined with the Silhouette method [23] applied to the original, non-graph representation of input data. This dataset includes all node features, along with additional information like the number of parking spots, existing charging stations in the area, and their power output in kilowatt-hours. After calculating the Silhouette index, we establish the optimal cluster number as $k = 3$, as this yielded a lower Silhouette index value compared to other potential values for k. On examining the cluster characteristics, we find a strong correlation between the cluster formation and the number of charging stations in each node area. Consequently, we select the candidate sites based on their potential: sites within clusters containing low number of existing charging stations are labelled as high potential, those with medium amount of charging stations as a medium potential, while clusters with highest number of stations, and hence the lowest need for new infrastructure, as low potential.

Table 1 – Battery capacity and range of the most sold EVs in the UK in year 2023 [3].

Vehicle Model	Battery Capacity [Kwh]	Range [km]
Tesla Model Y	60-75	455-542
MG4	51-77	349-520
Audi Q4 e-tron	82	455-543
Tesla Model 3	60-78	513-528
Polestar 2	82	555-653
Volkswagen ID.3	62-82	430-558
Kia e-Niro	68	463
BMW i4	83.9	413-589
Volkswagen ID.4	82	515-550
Skoda Enyaq	82	538-547

3.3 New infrastructure installation performance metric

To select the optimal changing location within a range of selected potential candidate sites, we focus on medium-high potential areas and opt to maximize the area demand coverage by simulating how the addition of a new charging station affects the incremental change in the area's total charging output, and propose Incremental Coverage Difference (ICD) metric. This metric is dependent on many factors and requires knowledge of: 1) total annual EV flow within the demand node, 2) average annual power requirement of an EV, and 3) maximal annual power output of a demand node. To measure the approximate number of EVs in the Glasgow City area, we multiply the number of registered private vehicles with the assumed 10% EV penetration rate:

$$N_{EV} = 0.1 \times 197{,}540 = 19{,}754 \tag{3}$$

In order to measure the charging frequency of an EV within city limits, we first calculate the average driving range of an EV. Given existing statistics on most purchased vehicles in the UK (see Table 1), the average driving range of an EV is set to $range_{avg}$ = 466.45 km, while average battery capacity is set to $capacity_{avg}$ = 68.20 kWh. According to the UK Department of Transport, total annual traffic in 2023 for private cars and taxis in Glasgow City is 2.684 billion kilometers. Thus, the average annual travel distance of cars in Glasgow is:

$$d_{CAR} = \frac{2{,}684.06 \times 10^6}{197{,}540} \frac{km}{year} = 13{,}567.4 \frac{km}{year} \tag{4}$$

Assuming that EVs only recharge when the battery capacity is lower than 20%, the yearly charging frequency of a car in the Glasgow City area is:

$$f_{EV} = 13{,}567.4 \frac{km}{year} \div \left(range_{avg} \times 0.8\right) = 36.36 \frac{charges}{year} \tag{5}$$

Similarly, the average energy need of an EV per year is:

$$p_{EV} = f_{EV} * \left(capacity_{avg} * 0.8\right) = 1{,}983.7 \frac{kWh}{year} \tag{6}$$

With this in mind, we calculate the charging capacity of each charging station within 500m radius. The 500m radius is chosen as a charging demand node radius as it is generally considered a comfortable walking distance to the charging station according to previous studies [24]. First, we calculate the annual EV flow within the area i, given 10% EV penetration rate:

$$flow_{EV}^{i} = 0.1 \times flow_{CAR}^{i} \tag{7}$$

Following this, we calculate the total annual energy requirement of a demand node:

$$C_{total}^{i} = flow_{EV}^{i} \times p_{EV} [kWh] \tag{8}$$

Next, we calculate the current annual power output of the demand node, given the average charger power output in the demand area $O_{avg} [kW]$, and maximal (24h) utilization:

$$C_{current}^{i} = 24 \times 365 \times no\_chargers^{i} \times O_{avg}^{i} [kWh] \tag{9}$$

Thus, current demand node coverage is measured as:

$$C^{i} = min\left[\frac{C_{current}^{i}}{C_{total}^{i}}, 1\right] [\%] \tag{10}$$

Similarly, to calculate the new demand node coverage when adding a charger with output power $c_{out}[kW]$ to a demand node i:

$$C_{new}^{i} = min\left[\frac{C_{current}^{i} + 24 \times 365 \times c_{out}}{C_{total}^{i}}, 1\right] [\%] \tag{11}$$

Lastly, to calculate the Incremental Coverage Difference (ICD), when adding a new charger to a demand node:

$$ICD^{i} = C_{new}^{i} - C^{i} [\%] \tag{12}$$

Figure 2 presents an algorithmic method for selecting optimal sites for charger installation based on their ICD scores for installation of chargers with a predefined power output. This process involves assessing the utility gain for demand nodes resulting from the addition of new charging stations in areas that are equipped with $n_{charging}$ chargers and $n_{parking}$ parking stations. High ICD scores are indicative of demand nodes with underserved charging needs due to insufficient infrastructure.

Conversely, sites meeting or surpassing demand are excluded from consideration. A key advantage of this approach is the ranking of candidate stations by their *ICD* scores, which offers a clear hierarchy of installation priority and anticipated impact on meeting the charging demand of individual charging nodes. Additionally, the algorithm addresses the practical scenario of bulk charger installation and considers how new charging infrastructure influences the charging capacity of nearby demand nodes. Additionally, it provides insights into the effects of different deployment strategies on charging demand, i.e., a comparison of utility of installing a number of *22kW* chargers compared to a number of *50kW* chargers within city limits. Finally, this solution advocates for leveraging existing parking spaces to identify ideal charging locations, thereby potentially reducing investment costs.

Figure 2 – Site Selection Algorithm for New Charging Stations

Algorithm 1: Site Selection Algorithm for New Charging Stations

Require: Charging nodes $\mathcal{V}$, Number of chargers to be installed k, New charger power output c_{out} [Kw], Charging node radius r

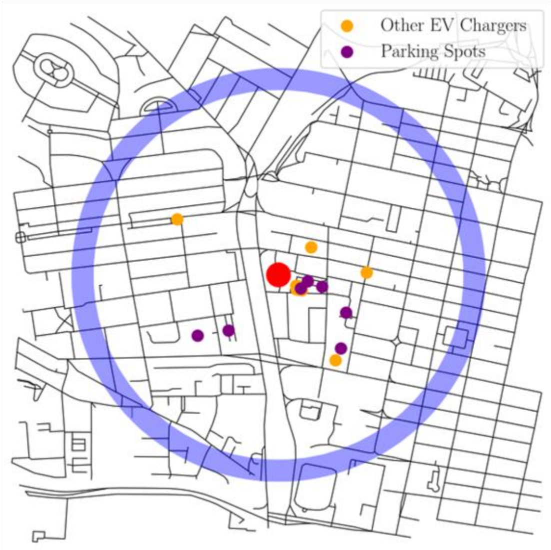
Initialize list of new stations: $newStations \leftarrow [\,]$
for $i = 1$ to k **do**
    $maxICD \leftarrow 0$
    $maxICDStation \leftarrow null$
    for each node $v \in V$ **do**
        Calculate ICD^v (Eq. 12)
        if $v_{type} == parking$ and $n^d_{parking} \geq 1$ **then**
            if $ICD^v > maxICD$ **then**
                $maxICD \leftarrow ICD^v$
                $maxICDStation \leftarrow v$
            end if
        end if
    end for
    Add site $maxICDStation$ to $newStations$
    for $v \in V$ **do**
        Calculate the distance $dist^v$ from node v to $maxICDStation$
        if $dist^v \leq r$ **then**
            $n^v_{parking} \leftarrow n^v_{parking} - 1$
            $n^v_{charging} \leftarrow n^v_{charging} + 1$
            Recalculate node output given newly added c_{out} (Eq. 11)
        end if
    end for
end for
return $newStations$

4 Data Collection

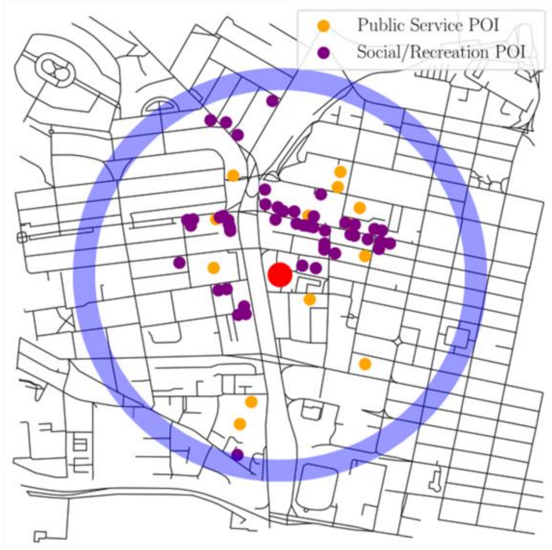
To comprehensively capture the diverse factors impacting charging demand, we build upon existing literature (for more details see [5]). To represent demand nodes, we consider an area in $r = 500m$ radius around a charging station or parking spot, aligning with the preference for shorter walking distances for car charging as noted in [24]. We incorporate a variety of node features, including Points of Interest (POI), population statistics, traffic flow, coverage, parking availability, and existing charging stations. Additionally, the study integrates aspects related to the Human Development Index, as detailed in the Scottish Index of Multiple Deprivation (SIMD) dataset [25].

Figure 3 visually depicts the construction of these charging demand nodes. Each subfigure highlights a demand node centered on an existing charging station or parking spot, with a 500-meter radius encompassing relevant metrics. It's noteworthy that SIMD data is organized into "data zones", which are specific areas designated for small-scale statistics in Scotland. Table 2 outlines the various data sources used in assembling datasets used in this study. OpenStreetMap [27] provides extensive details on street-level parking and POIs, including counts of recreational/touristic and social/public service POIs. This information is valuable for identifying locations with existing parking infrastructure suitable for new EV charging station installation, potentially reducing costs and expediting deployment. Additionally, the distribution of POIs aids in strategically placing charging stations in areas where drivers are likely to spend considerable time, enhancing charging convenience. Traffic flow data within Glasgow City is collected from [28], where only data for cars and taxis is selected. Existing charging infrastructure data is collected and pre-processed through the UK National Chargepoint Registry [26], while socio-demographic data is collected through the Scottish Index of Multiple Deprivation (SIMD) [25]. SIMD contributes in-depth socio-demographic data, encompassing income, employment, health, education, service access, and crime statistics. Incorporating SIMD data into the analysis offers valuable insights into the socioeconomic context of potential EV charging station sites, ensuring that infrastructure is optimally placed to be effective and beneficial, especially in areas that might otherwise lack adequate EV infrastructure.

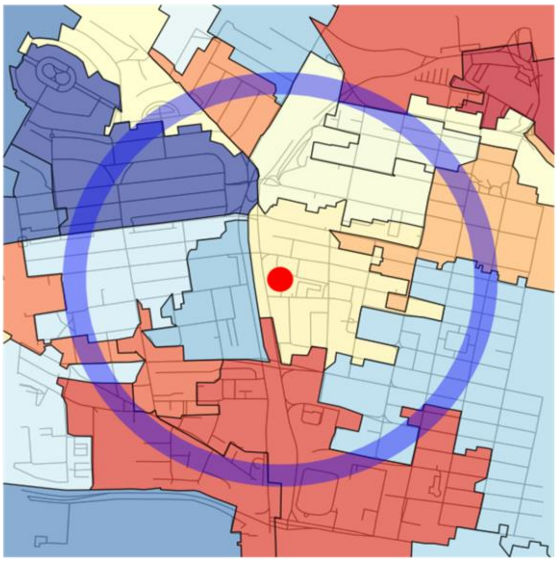
Figure 3 – Data sources collected and used in this study.



a) Charging and parking spots



b) Public service and social POI



c) Deprivation data (SIMD)

Table 2 – Experimental results of proposed method for k = 300 placed stations.

Dataset	Core Aspect	Selected Attributes
SIMD [25]	Socio Demographic Factors	total population, working population, shape area, income rate, crime rate, lat, long, avg. car drive to public service POIs, avg. public transport drive to public service POIs
UK National Chargepoint Registry [26]	Existing Charging Infrastructure	output power (kW), connector type, lat, long
OpenStreetMap [27]	POIs, Parking	lat, long, rec poi: {'restaurant', 'cafe', 'bar', 'theatre', 'arts centre', 'cinema'}, soc poi: {'school', 'community centre', 'music school', 'prep school', 'art school', 'university', 'hospital', 'kindergarden', 'pharmacy', 'place of worship', 'bank', 'doctors', 'dentist', 'kindergarten', 'social facility', 'post office', 'clinic', 'childcare', 'library', 'police', 'veterinary', 'nursing home', 'post depot', 'fire station', 'courthouse', 'nursery', 'healthcare', 'bank'}, parking: {'fuel', 'parking', 'events _ venue', 'college', 'hospital', 'university', 'venue', 'food court', 'exhibition centre', 'community centre', 'townhall', 'stadium', 'conference centre'}
Road Traffic Statistics, UK Department of Transport [28]	Traffic Flow	cars_and_taxis, lat, long

5 Experimental Evaluation

The experimental results of the proposed methodology are detailed in Table 3. To ensure a comprehensive evaluation, our ChargeDEM GNN-based approach is compared against four distinct algorithms: Kmeans clustering, Spectral clustering, Agglomerative clustering, and GraphSAGE [29]. K-means and spectral clustering analyze the data solely based on its inherent features. On the other hand, GraphSAGE employs an unsupervised graph representation technique, aggregating feature information from a node's location and environment. The superior ICD performance of the ChargeDEM model signifies its enhanced capability to integrate spatial and demographic data with EV charging characteristics, making it more adept at pinpointing strategic locations for urban charging station deployment. The improved ICD scores across different station output values highlight the nuanced considerations necessary for effective EV infrastructure planning.

Our model was particularly adept at identifying high-potential areas for new charging infrastructure, which is crucial for expanding the EV charging network efficiently and equitably. The use of GNNs enabled the model to capture complex spatial and relational data effectively. Additionally, the model takes into account how installation in one node affects the neighboring nodes, an often-overlooked aspect in heuristic-based approaches.

For the site selection task, we compare the ICD improvement over $k = 300$ installed stations for the four most common output power values: 7.1kW, 22kW, 34kW and 50kW. Figure 4 showcases the percentage of ICD improvement over the number of installed stations, providing a visual representation of how strategic deployment can maximize utility and justify investment. The similarity in utility gained from installing 60 charging stations at both 7.1kW and 50kW suggests that slow charging stations may be adequate in areas within Glasgow City limits, negating the need for higher output stations within city environments. This observation points towards a potentially lower need for rapid charging infrastructure within urban centers, which is often more expensive to install, leaving the conclusion that such investment should be highly targeted and potentially only required near highways. The 22kW stations emerge as a pragmatic choice, offering a middle ground that caters well to a wider range of EV charging needs, both fast and slow, while still providing considerable ICD improvement. This insight is invaluable for stakeholders, as it indicates that a balanced approach to station capacity could serve the diverse needs of EV users without incurring excessive installation costs.

Figure 4 – CD improvement per energy capacity for the Glasgow City area for k = 300 installed charging stations.

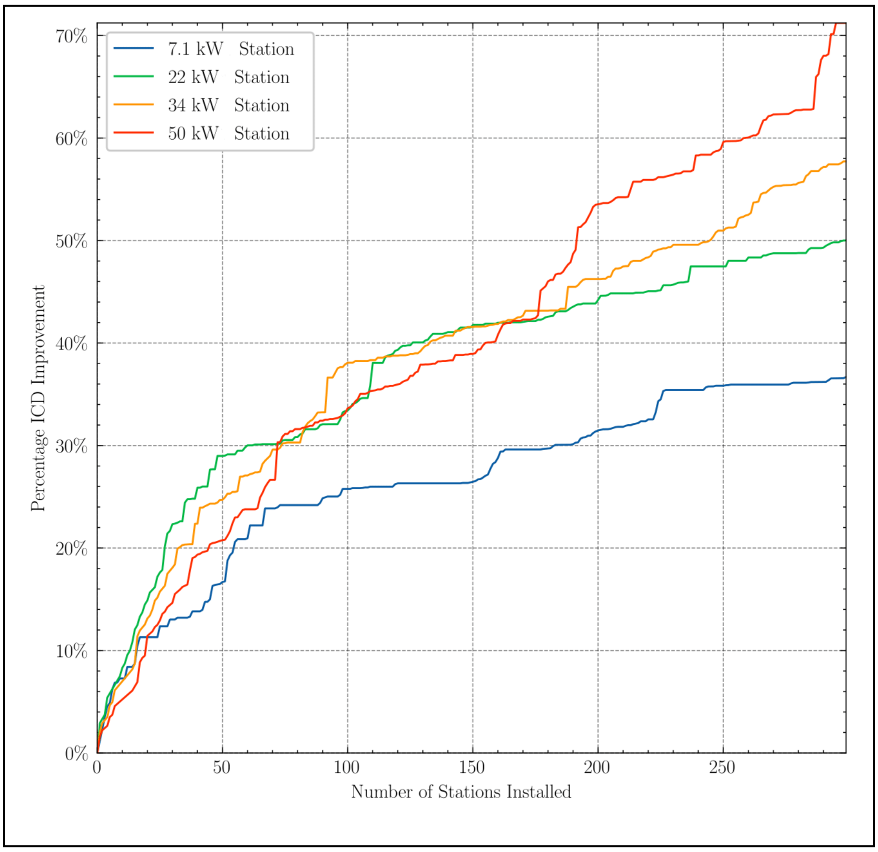


Table 3 – Experimental results of proposed method for k = 300 placed stations.

Algorithm	$ICD_{7.1kW}$	ICD_{22kW}	ICD_{34kW}	ICD_{50kW}
K-Means	0.2578	0.3756	0.5065	0.5718
Spectral	0.3374	0.4675	0.5376	0.5895
Agglomerative	0.3354	0.4684	0.5293	0.5875
GraphSage	0.2723	0.3893	0.5148	0.5544
ChargeDEM (Ours)	0.3669	0.5001	0.5772	0.7123

Additionally, as the figure shows, the incremental gains plateau as more stations are added, which could inform decision-makers about the point of diminishing returns and help optimize the number of stations to install. Furthermore, the results open up discussions about how different urban areas may require tailored strategies based on their unique demographic and spatial characteristics. The ChargeDEM model's adaptability to various urban settings could be explored in future research, potentially leading to a more customizable framework for infrastructure planning.

6 Conclusion and Future Work

This study provides a geodemographic-aware placement of EV charging stations through the usage of graph neural networks and quantification of the utility of new charging infrastructure. To address the limitations of existing site selection approaches for EV charging station placement, through this approach, we address three important factors: modeling of complex spatial relationships within the charging demand network, injection of demographic and human development information into the site selection process, and quantification of the impact of placed EV charger, both on the immediate charging demand node, as well as neighboring nodes. Experimental results performed on the case study of Glasgow City suggest that our methodology helps identify charging stations that provide the highest utility and leads to improved charging station utilization.

The proposed framework introduces three main challenges for future research. First, there’s a need to assess the specific spatial contexts of the charging demand nodes, such as distinguishing between residential, business, or high-traffic areas. This differentiation could enhance the clustering phase, potentially revealing distinct charging patterns linked to different area types. Second, the approach to graph construction could be reevaluated. Instead of relying solely on distance, alternative metrics like travel or commute time might yield more effective connections between nodes. Finally, the framework offers the possibility of being framed as a reinforcement learning problem. In this scenario, the utility function described in the current study would serve as the basis for the reward function, guiding the learning process towards more effective solutions.

Acknowledgment

This project has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 955422.

References

[1] T. Bruckner, I. A. Bashmakov, Y. Mulugetta, *et al.*, “Chapter 7 - energy systems,” in *Climate Change 2014: Mitigation of Climate Change. IPCC Working Group III Contribution to AR5*, Cambridge University Press, Nov. 2014.

[2] R. Peng, J. H. C. G. Tang, X. Yang, M. Meng, J. Zhang, and C. Zhuge, “Investigating the factors influencing the electric vehicle market share: A comparative study of the european union and united states,” *Applied Energy*, vol. 355, p. 122327, 2024.

[3] Society of Motor Manufacturers and Traders, *New uk ev and afv registrations*, [Online; accessed 21-December-2023], 2023.

[4] S. Haustein and A. F. Jensen, “Factors of electric vehicle adoption: A comparison of conventional and electric car users based on an extended theory of planned behavior,” *International Journal of Sustainable Transportation*, vol. 12, no. 7, pp. 484–496, 2018.

[5] L. Adenaw and S. Krapf, “Placing bev charging infrastructure: Influencing factors, metrics, and their influence on observed charger utilization,” *World Electric Vehicle Journal*, vol. 13, no. 4, p. 56, 2022.

[6] C. Bian, H. Li, F. Wallin, A. Avelin, L. Lin, and Z. Yu, “Finding the optimal location for public charging stations–a gis-based milp approach,” *Energy Procedia*, vol. 158, pp. 6582–6588, 2019.

[7] G. Battapothula, C. Yammani, and S. Maheswarapu, “Multi-objective simultaneous optimal planning of electrical vehicle fast charging stations and dgs in distribution system,” *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 4, pp. 923–934, 2019.

[8] D. Efthymiou, K. Chrysostomou, M. Morfoulaki, and G. Aifantopoulou, “Electric vehicles charging infrastructure location: A genetic algorithm approach,” *European Transport Research Review*, vol. 9, no. 2, pp. 1–9, 2017.

[9] J. Qin, J. Qiu, Y. Chen, T. Wu, and L. Xiang, “Charging stations selection using a graph convolutional network from geographic grid,” *Sustainability*, vol. 14, no. 24, p. 16797, 2022.

[10] B. Yıldız, O. Arslan, and O. E. Karas¸an, “A branch and price approach for routing and refueling station location model,” *European Journal of Operational Research*, vol. 248, no. 3, pp. 815–826, 2016.

[11] C. Upchurch and M. Kuby, “Comparing the p-median and flowrefueling models for locating alternative-fuel stations,” *Journal of Transport Geography*, vol. 18, no. 6, pp. 750–758, 2010.

[12] A. Vavouris, L. Stankovic, and V. Stankovic, “Integration of drivers’ routines into lifecycle assessment of electric vehicles,” in *8th International Electric Vehicle Conference*, 2023.

[13] C.-N. Wang, N.-A.-T. Nguyen, and T.-T. Dang, “Offshore wind power station (owps) site selection using a two-stage mcdm-based spherical fuzzy set approach,” *Scientific reports*, vol. 12, no. 1, p. 4260, 2022.

[14] A. C¸. Boyacı and A. S¸is¸man, “Pandemic hospital site selection: A gisbased mcdm approach employing pythagorean fuzzy sets,” *Environmental Science and Pollution Research*, vol. 29, no. 2, pp. 1985–1997, 2022.

[15] Y. Rezaeisabzevar, A. Bazargan, and B. Zohourian, “Landfill site selection using multi criteria decision making: Influential factors for comparing locations,” *Journal of Environmental Sciences*, vol. 93, pp. 170–184, 2020.

[16] T. Lan, H. Cheng, Y. Wang, and B. Wen, “Site selection via learning graph convolutional neural networks: A case study of singapore,” *Remote Sensing*, vol. 14, no. 15, p. 3579, 2022.

[17] B. Yu, H. Yin, and Z. Zhu, “Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting,” *arXiv preprint arXiv:1709.04875*, 2017.

[18] T. N. Kipf and M. Welling, “Variational graph auto-encoders,” *arXiv preprint arXiv:1611.07308*, 2016.

[19] J. Park, M. Lee, H. J. Chang, K. Lee, and J. Y. Choi, “Symmetric graph convolutional autoencoder for unsupervised graph representation learning,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6519–6528.

[20] P. Velickoviˇ c, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y.´ Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.

[21] H. Zeng, H. Zhou, A. Srivastava, R. Kannan, and V. Prasanna, “Graphsaint: Graph sampling based inductive learning method,” *arXiv preprint arXiv:1907.04931*, 2019.

[22] S. Lloyd, “Least squares quantization in pcm,” *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[23] P. J. Rousseeuw, “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,” *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[24] P. Van der Waerden, H. Timmermans, and M. de Bruin-Verhoeven, “Car drivers’ characteristics and the maximum walking distance between parking facility and final destination,” *Journal of transport and land use*, vol. 10, no. 1, pp. 1–11, 2017.

[25] Scottish Government, *Scottish index of multiple deprivation 2020*, Online, Accessed: 2024-01-04, 2020.

[26] Department for Transport, *Uk national charge point registry api documentation*, Accessed: 2024-01-05, 2024.

[27] OpenStreetMap contributors, *Planet dump retrieved from https://planet.osm.org*, https://www.openstreetmap.org, 2017.

[28] Department for Transport, *Road traffic statistics*, https://www.gov.uk/ government/collections/road-traffic-statistics, 2023.

[29] W. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” *Advances in neural information processing systems*, vol. 30, 2017.