Extending EV Battery Lifetime: Digital Phenotyping Approach for Departure Time Prediction

YONGGEON LEE, Hanyang University, Republic of Korea
WOOJIN SONG, KENTECH, Republic of Korea
JUHYUN SONG*, KENTECH, Republic of Korea
YOUNGTAE NOH*, Hanyang University, Republic of Korea

Battery degradation, a gradual loss of usable capacity over time, is one of the major hurdles for widespread adoption of electric vehicles (EVs). We introduce *delayed full-charging (DFC)* algorithm to mitigate degradation and extend the lifetime of EV batteries in battery management systems (BMS). When the EV is plugged in, the *DFC* algorithm charges batteries up to approximately 80% state of charge (SOC) and delays full charging until the predicted unplug time (t_{unplug}). This approach significantly reduces the time batteries remain fully charged ($t_{100\%}$), thereby mitigating degradation while ensuring charging time for EV users to utilize the full battery capacity. For predicting t_{unplug}, we propose a novel methodology that uses digital phenotyping to predict departure times. This method leverages smartphone data to capture irregular but predictable departure patterns by reflecting relevant behavioral and environmental contexts. A case study with 48 participants was conducted to empirically evaluate the departure time prediction performance using tree-based ensemble models trained on smartphone data, compared to a baseline Long Short-Term Memory (LSTM) model trained on historical data. Results reveal that models utilizing mobile passive features achieved a Mean Absolute Error (MAE) as low as 2.18 hours on weekdays and 4.46 hours on weekends, demonstrating superior effectiveness in capturing irregular patterns compared to the baseline model trained only on historical temporal features.

CCS Concepts: • **Human-centered computing → Empirical studies in ubiquitous and mobile computing**.

Additional Key Words and Phrases: Delayed Full-Charging (DFC), Battery Management System (BMS), Departure Time Prediction, Digital Phenotyping

ACM Reference Format:
Yonggeon Lee, Woojin Song, Juhyun Song, and Youngtae Noh. 2024. Extending EV Battery Lifetime: Digital Phenotyping Approach for Departure Time Prediction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 4, Article 201 (December 2024), 30 pages. https://doi.org/10.1145/3699725

1 Introduction

In recent years, advancements in lithium-ion battery (LIB) technology and the rapid expansion of the electric vehicle (EV) market [15] have paved the way for innovative eco-friendly transportation solutions to combat climate change. However, these developments have also brought the critical issue of battery degradation to the forefront. Battery degradation entails the progressive reduction in capacity and performance over time, leading to reduced lifespan and acceleration capability of EVs [3, 44]. Thus, mitigating battery degradation has significant implications for the widespread adoption and viability of EVs in the era of carbon neutrality [20].

*Junyun Song and Youngtae Noh are the corresponding authors.

Authors' Contact Information: Yonggeon Lee, yonggeonlee@hanyang.ac.kr, Hanyang University, Seoul, Republic of Korea; Woojin Song, wsong@kentech.ac.kr, KENTECH, Naju-si, Republic of Korea; Juhyun Song, jsong@kentech.ac.kr, KENTECH, Naju-si, Republic of Korea; Youngtae Noh, youngtaenoh@hanyang.ac.kr, Hanyang University, Seoul, Republic of Korea.

ACM 2474-9567/2024/12-ART201
https://doi.org/10.1145/3699725

Battery degradation poses significant challenges for both EV users and manufacturers. For users, degradation results in reduced battery capacity, leading to a limited driving range and more frequent charging. It also increases maintenance costs, as users may need to replace the batteries or buy a warranty on the batteries. On the other hand, manufacturers must meet user demand while complying with strict environmental regulations, where battery lifespan is a crucial metric. For example, the EU mandates proper documentation of battery lifespan for EVs starting in 2024, with minimum values to be set by 2027. More countries are expected to adopt similar regulations on batteries and EVs soon. By 2030, it is anticipated that over 5 million metric tons of LIBs will reach the end of their lifespan [5], emphasizing the need for effective technologies to mitigate battery degradation.

In fact, most EV users inadvertently speed up battery degradation by regularly charging their vehicles overnight at home or for long periods at work. This increases the duration of batteries remain at 100% state of charge (SOC), denoted as $t_{100\%}$, a condition where battery aging accelerates significantly [29]. At 100% SOC, batteries are more prone to parasitic chemical reactions such as growth of passivation layers[1] and dissolution of transition metals[2] compared to lower SOC. These processes consume cyclable lithium and chemically damage active materials, leading to battery degradation [13, 28]. Therefore, the longer the battery remains at a high SOC, the faster it degrades. To mitigate the adverse effects of prolonged $t_{100\%}$ on battery degradation, a direct method is to manually limit the upper SOC for as long as possible through battery management system (BMS). Several studies have addressed the concept of delayed charging, which prevents the EV from reaching a fully charged state [22–24, 33, 35, 38, 39]. However, previous delayed charging schemes involve a simple process of pausing and recharging, which relies on user involvement or specified values to decide when to resume charging. A more effective approach would be to automate the charging process using data-driven algorithms.

To address this issue, we propose *delayed full-charging* (*DFC*) algorithm to mitigate the degradation by active management of $t_{100\%}$. The *DFC* algorithm postpones full charging and then resumes fast charging to reach 100% SOC just before the EV is expected to start driving. By delaying full charging, *DFC* can significantly reduce $t_{100\%}$ and mitigate battery degradation. At the same time, it is crucial to accurately predict the end of the charging (unplug time, t_{unplug}), ensuring that the vehicle is fully charged right before a trip. This minimizes the risk of EV leaving earlier than predicted with partially charged battery, which could restrict travel distance. t_{unplug} is determined by the EV user's departure time, which can be analyzed through human behavior modeling. Mobile devices can capture diverse aspects of human behavior, uncover social patterns in daily activity, infer relationships, and identify socially significant places [11]. By leveraging human contextual data collected from smartphone, we expect to quantify an individual's pre-departure behavior. To achieve this, we propose a new methodology for predicting daily departure time through digital phenotyping [55]. To the best of our knowledge, our work represents the first attempt to use human data for predicting departure time by quantifying pre-departure behavior. Previous studies [8, 26, 45, 59] have investigated the use of smartphone GPS data to analyze mobility traces of individuals for predicting departure time. However, these approaches are limited as they primarily focus on identifying similar movement patterns in historical trajectories or incorporating time-related information, rather than leveraging diverse contextual information about human behavior.

One common concern among EV users is the miles per full charge as measured by the Worldwide Harmonized Light Vehicles Test Procedure (WLTP) [12]. That is, EV users tend to maintain a high SOC to ensure sufficient travel distance, often results in vehciles being frequently left fully charged for extended periods, exposing them to prolonged $t_{100\%}$. This is where the *DFC* algorithm comes into play. The motivation behind the *DFC* algorithm

[1]The passivation layer at the negative electrode (anode), commonly known as the solid electrolyte interphase (SEI), forms immediately upon contact between the metal and electrolyte. While it serves as a barrier with high electronic resistivity, which is crucial for battery safety, performance, and longevity, continuous SEI growth results in degradation. For more detailed information, see [4].

[2]Transition metal ions (such as nickel, manganese, or cobalt) are gradually oxidized in the idle state, releasing electrons and dissolving into the electrolyte from the positive electrode (cathode). These ions can react with the electrolyte and further contribute to SEI growth. For more detailed information, see [7].

is to mitigate the negative impact of unnecessarily prolonged 100% SOC on battery degradation by minimizing $t_{100\%}$. Thus, the *DFC* algorithm is applicable to scenarios where EVs remain at full charge for extended periods, such as when users charge their vehicles overnight at home—a common daily routine. This work specifically targets the scenario where users charge their EVs overnight at home and depart the next day. On weekdays, users are likely to charge and use their EVs for regular commuting between home and the workplace (*e.g.*, office, store, laboratory, etc). On weekends, however, they might use their fully charged EVs for various appointments or activities, following less predictable schedules. By examining this target scenario, we can smoothly extend our findings to similar situations, enhancing the *DFC* algorithm's generalizability. In addition, we can extend our work to various fields that leverage personalized mobility patterns of EV users, by applying departure time prediction method. For instance, our method can be utilized for thermal management, such as preheating the battery before departure to improve battery efficiency [19].

Finally, we emphasize the anticipated effects of the *DFC* algorithm on improving battery lifespan, drawing on findings from previous studies. In [29], battery degradation was evaluated over 10 months of storage at various SOCs and temperatures, showing significant capacity fade at 100% SOC. This fade was notably more pronounced compared to 90% SOC, with about a 7% greater reduction in relative capacity. Similarly, [36] examined the evolution of calendar aging over 24 months under comparable conditions. The results demonstrated a capacity fade gap of up to 15% between 90% and 100% SOC, with the disparity increasing at higher temperatures. These experimental results from previously well-established studies support the *DFC* algorithm's potential to mitigate battery degradation. Our ultimate future goal is to evaluate its effectiveness in extending battery lifespan through both laboratory and real-world deployments. The contributions of *DFC* are four-fold, as outlined below:

- We design an algorithm that mitigates the battery degradation by delaying full charging to reduce storage periods of high SOC, addressing common, but indifferent behavior of EV users who charge their vehicles overnight at home.
- We propose a novel approach that leverages digital phenotyping for departure time prediction by using smartphone data to capture relevant behavioral and environmental contexts. This enables an accurate, real-time decision-making process, ensuring the practicality of the algorithm.
- To address variability in departure patterns across different days and achieve personalization, we employ a DBSCAN-based outlier detection method with kneedle approach. This method automatically adjusts *epsilon* and *MinPts* parameters based on daily individual contexts, significantly enhancing adaptability to new observations.
- We conduct a user study involving 48 participants to evaluate our digital phenotyping-based departure time prediction method on a real-world dataset collected through our self-developed application. We identify relevant mobile passive features correlated with departure time and validate the predictive performance.

2 Related Work

Extensive research has been conducted to develop charging algorithms aimed at extending the lifespan of EV batteries. These algorithms adopt delayed charging focusing on $t_{100\%}$ based on the simulation results indicating that period at high SOC significantly accelerates the battery degradation [10, 24, 33, 50]. For the practical implementation of delayed charging, determining the optimal time to resume charging is essential. A substantial body of research has addressed this key parameter in the field of EV charging behavior, referring to it in various terms such as departure time, charging end time, or disconnection time [6, 16, 17, 34]. On the other hand, numerous studies have analyzed departure time in the context of mobility patterns, which are inherently linked to individual behaviors [8, 26, 45]. Accordingly, we summarize existing work on delayed charging (Section 2.1), and two related fields addressing departure time prediction: EV charging behavior prediction (Section 2.2) and smartphone behavior tracking (Section 2.3).

2.1 Delayed Charging

To improve battery lifespan, numerous optimized battery charging methods have recently been developed for real-world deployments in the smart device market. These methods adopt delayed charging approach to mitigate battery degradation by limiting the time spent at full charge, thus optimizing overall battery health and performance using BMS. Smartphone manufacturers like Apple and Samsung have integrated smart charging features (Optimized Battery Charging and Adaptive Battery) into their latest devices [2, 48]. These features learn users' daily charging habits, such as sleep and usage patterns, pausing charging at 80% or 85% and resuming just before the expected usage time. Similar implementations have been developed in the EV BMS sector, with initial work demonstrating effectiveness in improving battery life and reducing total costs based on simulation results [10, 42, 50, 62].

Hoke *et al.* [22] developed an intelligent charging method that optimizes both the costs of electrical energy and battery degradation. They proposed and compared various charging schemes, including: 1) charging at midnight, 2) a slow (6.6 kW) full charging before the vehicle leaves, 3) a fast (18 kw) full charging 30 minute before departure. They reported that extensive simulations on the proposed delayed charging methods can extend battery life by up to 50% compared to other schemes. In [24], similar results were obtained through extended simulations using experimental plug-in hybrid EV data from a controlled environment to reflect real-world conditions. Subsequently, the researchers introduced a partially nightly charging method that uses predicted energy needs for the next day (next day's vehicle usage) to charge the battery as required. Their simulation results demonstrated that this method can extend battery life by up to 150% over uncontrolled charging with a strong assumption that the next day's energy requirements can be perfectly predicted [23]. Lacey *et al.* [33] identified a linear relationship between average SOC and capacity loss in their studies analyzing factors influencing battery degradation. Their delayed charging scheme, which starts slow charging (3 kW) at an ideal time to achieve full charge just before departure, resulted in an 8.95% increase in battery lifespan due to a 10% decrease in average SOC over five months of laboratory tests. More recently, Montes *et al.* [38] designed two degradation-conscious charging schemes that take SOC into account through models that: 1) delay the charging until the end of the charging to reach the desired SOC (delayed smart model) and 2) balance the impact of high charging rate and high SOC (C-rate & SOC model). These models were found to reduce total operation costs by around 15% compared to immediate charging, through simulation tests over 384 unique scenarios with different combinations of SOC and diverse degradation factors.

Previous work has highlighted the significance of period at high SOC as a key battery degradation factor through delayed charging schemes. However, these studies did not provide a holistic view of practical charging mechanisms, nor did they explain the pausing and recharging process in detail. More importantly, the critical parameter for determining when to resume charging—departure time-is limited, as it has typically been set either deterministically (*e.g.*, specific time or user input) or statistically designed for scheduling objectives. This approach requires consistent user involvement and uncertainty regarding departure time, which can lead to user dissatisfaction. As a first step towards addressing this issue, we propose real-time departure time prediction method using digital phenotyping to enable the practical implementation of the *DFC* algorithm. This will be discussed in detail in Section 3.2.

2.2 EV Charging Behavior Prediction Using Historical Data

In the context of advancements in EV charging methods, extensive research has been conducted to understand EV charging behaviors [16, 25, 51, 52]. These behaviors include departure and arrival times, charging duration, and other relevant charging patterns derived from EV drivers. Understanding these patterns is crucial for designing smart grid modernization that integrates EV charging infrastructure, such as infrastructure planning, resource allocation, consumer satisfaction, and the profitability of charging stations [65]. To achieve this, recent studies have leveraged historical data and adopted various machine learning approaches to predict departure times based on the temporal patterns [51, 65].

Xu [64] used a Support Vector Machine (SVM) to forecast EV departure times at UCSD campus to enhance power grid management efficiency. The model was trained on historical data of departure times as well as temporal features (*e.g.*, day of the week and hours) achieving an average Mean Absolute Percentage Error (MAPE) of 3.7%. However, the study's scope was very limited as it focused on weekday commuting within a specific campus, which may not be generalizable to other scenarios. Frendo *et al.* [17] predicted EV departures using regression models trained on historical data and vehicle information to improve allocation of charging resources among multiple vehicles by charging scheduling. Among these models, XGBoost performed the best and historical departure time was identified as the most important feature. Khwaja *et al.* [30] used a Long Short-Term Memory (LSTM) network which is effective for sequential data, to forecast charging start time and duration in the residential sector. Despite this, they oversimplified the task by discretizing target values into 24 levels due to high variability. Similarly, Boulakhbar *et al.* [6] utilized LSTM approach to predict EV departure times at charging stations in Morocco to enhance EV charging management. This work was also yet limited by its constrained scenario of focusing solely on e-taxi drivers charging at public stations. In [34], online machine learning models were proposed to predict daily EV departure times during the COVID-19 pandemic using historical features of charging sessions and trips. Although these models demonstrated the adaptability to changing departure patterns between pre-pandemic and post-pandemic periods by achieving a best mean absolute error (MAE) of 2.75 hour, they failed to reduce the prediction boundary which exceeded 9 hours.

As described above, previous studies on EV departure time prediction have primarily relied on historical data, which demonstrated reasonable performance only in limited scenarios. This is because, under specific conditions such as time (*e.g.*, weekdays), demographics (*e.g.*, e-taxi drivers), and location (*e.g.*, campus), EV departures often exhibit regular and consistent patterns that can be captured by the temporal patterns inherent in historical data. However, EV departures can occur irregularly in diverse situations, such as weekday appointments and weekend trips in daily lives. In this work, we address the initial problem of capturing the irregularity of departure patterns. To this end, we re-examine the concepts of regularity and irregularity, as defined in the Cambridge Dictionary. Regularity is characterized as the consistent occurrence of events in a fixed pattern, with uniform or similar intervals between successive instances, whereas irregularity refers to events that do not occur at regular times. To illustrate this intuitively, we depicted areas of regularity and irregularity, along with a gray overlapping zone between them as shown in Figure 1(a). This zone represents departures that are irregular in nature but often obscured by individuals' routines, distinguishing them from the complete randomness associated with irregularity. Our objective is to distinguish between irregular, yet predictable departure patterns and complete randomness. We suggest that these irregular patterns can be inferred from both behavioral and environmental contexts relevant to departure. Figure 1(b) illustrates an example of the departure time predictions throughout the day, comparing the use of historical features alone with the inclusion of behavioral and environmental features. The historical-based approach detects regular departures at 9 a.m., while the combined approach accurately identifies all departure times, successfully reducing false positives that occur when using only historical data. Accordingly, we propose an alternative approach that leverages mobile passive data through digital phenotyping to capture behavioral and environmental contexts. As depicted in Figure 1, this approach surpasses the limitations of traditional methods focused on regular, temporal patterns, particularly evident during weekdays. By doing so, we significantly improve the decision process for resuming charging in the *DFC* algorithm, thereby better mitigating degradation and ensuring trip distance.

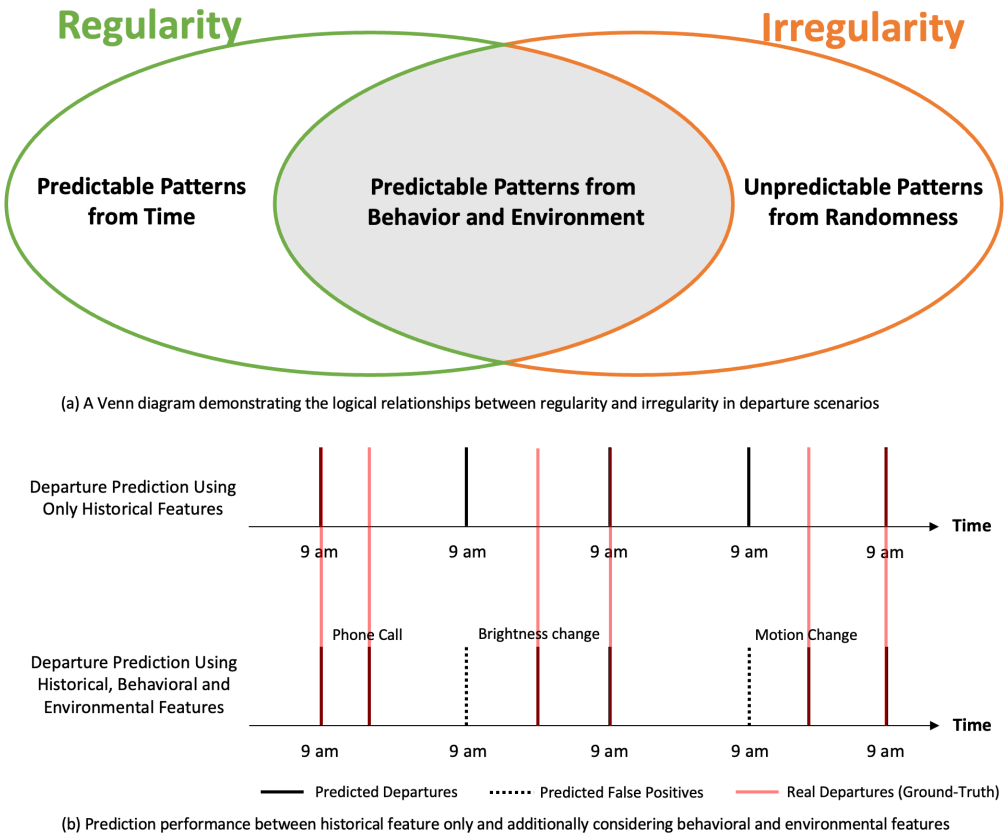


Fig. 1. (a) A Venn diagram illustrates the relationships between regular and irregular departure scenarios. (b) Predictive performances using historical features alone is compared against a model that includes behavioral and environmental features. Incorporating passive smartphone sensing data enhances accuracy by better identifying missed departures and reducing false positives, utilizing a blend of temporal, behavioral, and environmental data.

2.3 Smartphone-Based Behavior Tracking for Mobility Analysis

The advent of ubiquitous computing technologies has enabled the use of smartphones for human behavior analysis by unobtrusively collecting passive mobile data in daily life. In this regard, digital phenotyping provides diverse contextual insights into individuals through lifelogging data, which can significantly improve the performance of departure time prediction. Nevertheless, existing work has primarily relied on GPS data from smartphones for behavior tracking in mobility analysis. This approach has been used to predict mobility traces in real-world environments for diverse applications, including traffic control, trajectory recommendations, vehicle tracking, and travel mode detection [9, 18, 21, 27, 32, 60, 66].

Imai *et al.* [26] developed an early destination prediction model that incorporates trajectory tracking with next place prediction (NPP) by leveraging user contexts such as time of day and day of the week. This model achieved an accuracy of over 0.75 for predicting the first destination of the day, but its reliability was limited by the destination candidates that were technically derived from DBSCAN clustering. Addressing the challenge of tracking detailed trajectories, Sadri *et al.* [45] predicted continuous trajectories including departure times using

Dynamic Time Warping (DTW) on a distance matrix computed from positive correlations between historical trajectories (*e.g.*, morning and afternoon) across the same days of different weeks. Their experimental results showed a MAE of 2.8-2.9 hours for departure time predictions after 12 pm, but with high variation due to low adaptability to new routines. Wang *et al.* [59] proposed a time-variant Markov model based on Gibbs sampling to enhance adaptability to sparse trajectories and distinct mobility patterns at different times of the day, which traditional Markov models struggle to handle. However, they applied the same time partitions for the Markov chains across all individuals, despite the fact that individual mobility patterns may persist for varying lengths of time. Recently, mobility modeling using smartphone GPS has expanded to address travel mode detection. Zhou *et al.* [67] developed a smartphone-based person travel survey system using a hybrid approach combining random forest and merging algorithm to refine the inconsistency in travel mode predictions. However, this algorithm was somewhat intrusive, as it required individuals to verify the algorithm's predictions to establish ground truth. In [37], A large-scale mobility study involving 21,571 Android and iOS users over 8 weeks was conducted to analyze travel modes as part of a nationwide randomized controlled trial of transport pricing in Switzerland. The study found that travel mode detection faced challenges in distinguishing activities with similar speeds and in managing trip chaining.

With the widespread adoption of GPS-equipped smartphones, previous research has demonstrated the feasibility of using smartphones to track human mobility behavior. However, most studies have primarily focused on modeling mobility behavior using only GPS data or with limited contextual information obtrusively collected from surveys. Our approach exploits rich contextual information of an individual's environment and behavior through passive sensing data from smartphones. By doing so, we provide more accurate departure time predictions by identifying the inherent relationships between departure patterns and digital behavioral markers extracted from passive sensing data through digital phenotyping.

3 Delayed Full-Charging (DFC) Algorithm

We now introduce the *DFC* algorithm, detailing the charging schedule based on predicted departure times. To facilitate an intuitive understanding of its charging mechanism, we provide examples of the algorithm's application in virtual scenarios, highlighting the importance of accurate departure time prediction. Next, we propose our departure time prediction method, which learns pre-departure behavioral and environmental indicators from smartphone data, and explain the rationale for adopting a digital phenotyping approach.

3.1 DFC Algorithm in Charging Mechanism

The *DFC* algorithm follows the concept of delayed charging as discussed in Section 2.1, providing an innovative solution to mitigate battery degradation by avoiding unnecessarily long $t_{100\%}$. Instead of charging the battery to its full capacity immediately, the *DFC* algorithm delays full charging and then resumes fast charging to reach 100% SOC just before the EV's anticipated departure. This is achieved by predicting accurate t_{unplug}, the departure time. Based on the predicted t_{unplug}, charging is paused when the battery reaches SOC around 80%, a sweet spot between the onset of rapid degradation and ensuring sufficient trip distance [29]. Accordingly, the *DFC* algorithm relies on two key methods: 1) intelligent charging strategy for pausing and resume charging, and 2) an accurate departure time prediction method to determine when to resume charging. We first elaborate on the charging mechanism of the *DFC* algorithm, and then describe the digital phenotyping approach for departure time prediction in Section 3.2.

The *DFC* algorithm operates as follows. Once the EV is plugged in, the algorithm continuously monitors for a potential departure. Until a departure is detected, the algorithm pauses charging if the SOC is higher than 80% and initiates slow charging if the SOC is lower than 80%. To ensure the EV has sufficient time for full charging, the algorithm predicts a time 30 minutes before the actual departure. Additionally, it incorporates a time buffer

(t_{margin}) before this predicted pre-departure time to safeguard against late predictions, guaranteeing the EV is fully charged on time. However, an excessively large t_{margin} could reduce the algorithm's effectiveness in mitigating battery degradation. Therefore, the *DFC* algorithm is optimized to minimize degradation while maintaining a low probability of incorrect predictions (10 parts per million, ppm) through a selective coupling-based degradation model [53]. Finally, when the pre-departure time (*i.e.*, 30 minutes plus t_{margin} before departure) is predicted, the algorithm initiates fast charging to achieve 100% SOC. The overall process of this charging mechanism is detailed in Algorithm 1.

Algorithm 1 *DFC* Algorithm

Require: Current SOC *soc*, Extra time buffer t_{margin}

Define: Pre-departure time as 0.5 hour plus t_{margin} prior to the departure moment
Initialization: Start real-time pre-departure detection when EV is plugged in
while pre-departure time is not predicted **do**
    if $soc < 80\%$ **then**
        Proceed with slow charging
    else
        Pause charging
    end if
end while
Initiate fast charging

Figure 2 illustrates examples of charging scenarios derived by implementation of the *DFC* algorithm. In Figure 2(a), the successful implementation of the *DFC* algorithm is demonstrated, accurately learning the charging patterns during night and daytime. The algorithm effectively reduces unnecessary $t_{100\%}$ during repetitive charging cycles and avoids restrictions on driving range by performing a last-minute full charge before departure. However, as shown in Figure 2(b), if t_{unplug} is predicted to be later than the actual time, the battery may not be fully charged when the EV is needed, limiting the driving range. This undesired outcome caused by inaccurate departure time predictions must be minimized to prevent user dissatisfaction. The *DFC* algorithm strikes a balance between mitigating degradation due to prolonged $t_{100\%}$ and minimizing such undesired predictions. By achieving this balance, the *DFC* algorithm extends battery lifespan while ensuring the EV's driving range.

3.2 Departure Time Prediction: Digital Phenotyping Approach

To accurately predict departure time, we propose a digital phenotyping apparoach. Digital phenotyping is defined as the "moment-by-moment quantification of the individual-level human phenotype in situ using data from smartphones and other personal digital devices" [55]. While originally focused on identifying connections between disease subtypes and genetic variations, digital phenotyping can be broadly applied to model diverse human behaviors using smartphone data. Therefore, we leverage digital phenotyping to predict an individual's departure by quantifying the relevant context preceding departure. To effectively capture this pre-departure contextual information, we consider two essential aspects: behavioral and environmental indicators. We define three key behavioral indicators that imply individual routines and signals for departure, along with two environmental indicators to capture the individual's sleep and wake-up contexts as outlined below.

(1) ***Physical activities***. Actions such as motion, steps, or walking can reveal movement patterns associated with an individual's behavior before departure, such as getting out of bed in the morning.
(2) ***Social interactions***. Interactions on social media or phone calls often indicate typical behaviors of individuals preparing to leave for an appointment.

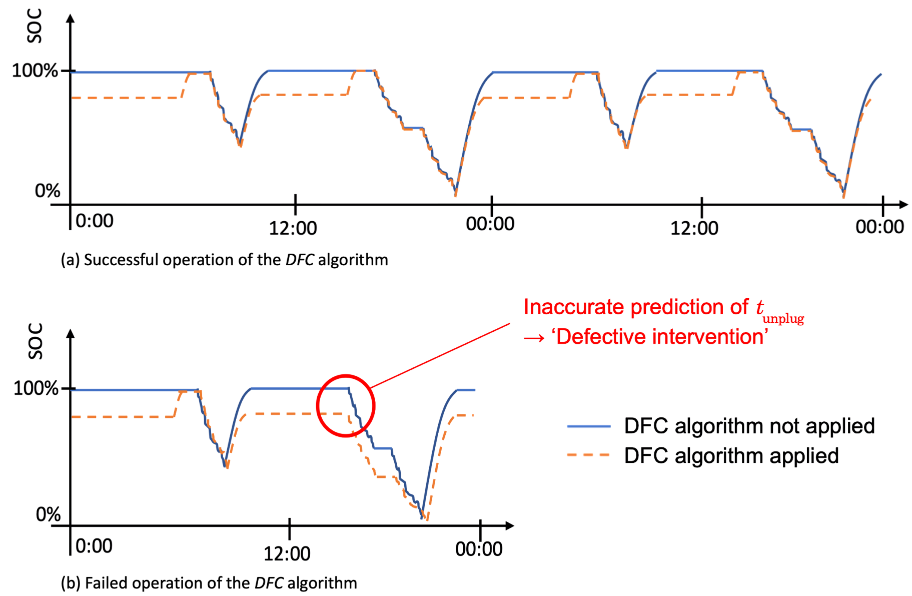


Fig. 2. Examples of virtual application of the *DFC* Algorithm. (a) Successful operation of the *DFC* algorithm: the algorithm effectively reduces the unnecessary $t_{100\%}$ when charging overnight or at work, and performs a full charge just before t_{unplug}. (b) Failed operation of the *DFC* algorithm: EV is not completely charged before departure, limiting its subsequent driving range due to the inaccurate prediction of t_{unplug} (defective intervention) around 4 p.m.

(3) ***Smartphone usage***. Smartphone activities like unlocking the phone or checking certain apps can also indicate an upcoming departure. For instance, individuals may use a navigation app to search for their destination or set alarms for their commute.
(4) ***Ambient light***. Changes in ambient light, such as increased brightness from opening curtains or turning on lights, imply that the individual has woken up.
(5) ***Ambient sound***. Changes in ambient sound, such as noises from activities like washing, preparing food, or dressing, can provide additional information about the individual's preparation for departure.

These 5 categories of indicators—physical activities, social interactions, smartphone usage, ambient light, and sound—can be passively collected from smartphones using API services and embedded sensors. By leveraging these behavioral and environmental indicators, our digital phenotyping approach provides comprehensive insights into individual departure routines associated with leaving home. By accurately predicting departure times, the *DFC* successfully reduce battery degradation and ensures sufficient trip distance. The schematic diagram of our departure time prediction is shown in Figure 3. Our method utilizes 9 types of sensors that represent the 5 categories of behavioral and environmental indicators, which will be discussed in Section 4.2. The detailed feature extraction process, including feature engineering and sliding window approach for real-time departure time prediction, will be described in Section 5.2.

4 Field Study and Dataset

In this section, we describe our large-scale field study, which strictly adhered to privacy protection guidelines under Institutional Review Board (IRB) approval. First, we outline the overall participant selection and ethics

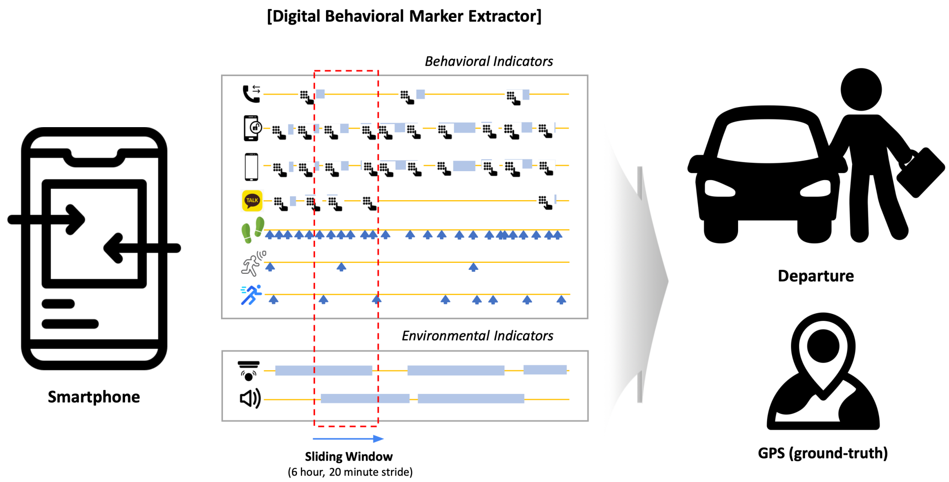


Fig. 3. Schematic diagram of EV departure time prediction using passive sensing.

process (Section 4.1). Next, we introduce our self-developed mobile client application for robust passive sensing data collection (Section 4.2).

4.1 User Study

This study aims to confirm the feasibility of exploiting digital phenotyping to enhance predicting departure behavior of EV users. Given the difficulty of recruiting EV users due to their low proportion, we expanded our target group to include individuals aged 18 to 69, regardless of gender (male: 234, female: 268, age mean = 28.5, age std = 10.6). Our preliminary work is based on the assumption that departure patterns would not differ significantly between EV users and non-EV users. More details on this assumption are presented in Section 8.4. The participants were recruited through various channels, including social networking sites (Facebook, Instagram), multiple universities, online communities, and poster advertisements in subways, targeting residents of Seoul and the metropolitan area in South Korea (Republic of Korea). The recruitment process occurred in three stages, each lasting between 2 to 3 months, spanning from May 31, 2021, to June 27, 2022, with observations concluding on July 5, 2022.

For non-obtrusive, in-the-wild data collection, participants were asked to install a self-developed mobile data collection application, EVA, which will be discussed in detail in Section 4.2. As a result, 506 Android users installed the application. They were provided with a comprehensive consent form with a detailed description of the study, including its purpose, procedures, and potential risks and benefits. To acknowledge their participation, they received up to $50 based on their engagement, compensating them for the use of personal data and encouraging participation. For privacy protection, all collected data were encrypted using AES-256-CBC and stored on a secure Linux-based server, ensuring participant privacy through anonymization. Data are retained for up to five years and are not shared outside the research team or used for commercial purposes. The dataset strictly adheres to ethical guidelines established by our IRB.

4.2 Data Collection

We developed an Android mobile application, EVA (EV Analyzer) for automatic passive sensing data collection. EVA continuously collects data in the background, aggregates it into small packets at regular intervals and

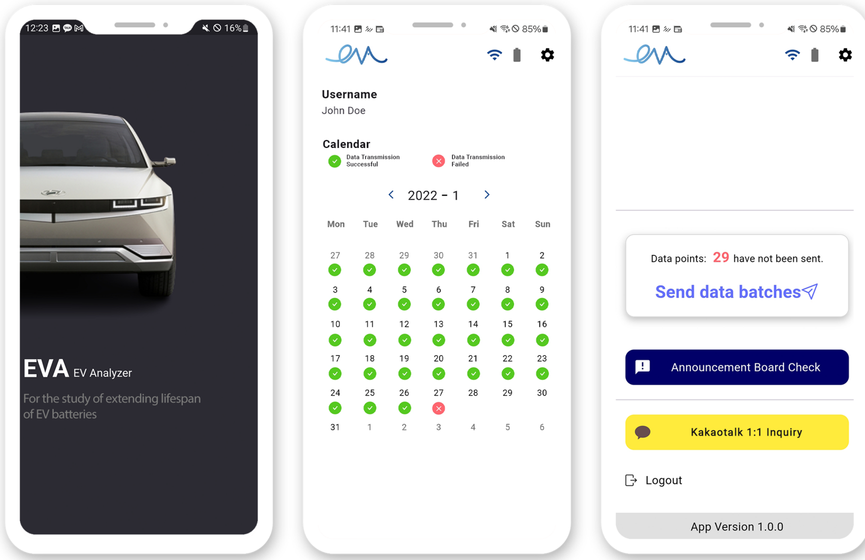


Fig. 4. Screen shots of EVA: splash screen, calendar interface, and manual data-uploading function.

transmits them to the cloud server of the data platform [56] when the smartphone is connected to Wi-Fi or a cellular network (*e.g.*, LTE). To ensure data quality, EVA provides data monitoring and management features for users. The home screen of EVA includes a calendar interface that marks the daily status of successful data collection. It also displays the amount of data stored on the smartphone and offers a data-uploading function, allowing participants to manually send data to the server if automatic transfer fails. This enables them to easily monitor and manage their data to maintain high data quality. Figure 4 illustrates sample screenshots of EVA splash screen, calendar interface, and data-uploading function.

Passive sensing data. As briefly introduced in Section 3.2, EVA collects data from 9 different passive sources (*i.e.*, activity type, steps, significant motion, calls, screen state, unlock state, app usage, ambient light and sound) which represent behavioral and environmental indicators for predicting departure time, along with GPS data for ground truth, totaling 10 types of passive data. These data sources are categorized as either continuous sensing-based or event-based, depending on their sampling method. Continuous sensing-based data were collected at regular intervals (*e.g.*, every 15 minutes), which was determined to be the optimal balance between sample resolution and battery consumption. Event-based data were recorded when triggered by specific events such as changes in phone state, physical activity transitions, or application usage. Table 1 provides detailed descriptions and sampling rates for each sensor source. EVA collects these sensor data running in the background, without requiring any user intervention.

Ground-truth data. To obtain ground-truth data, we extracted departure times from GPS samples, which provide longitude and latitude coordinates at intervals between 5 to 15 minutes. First, we identified distinct location labels using Point of Interest (POI) clustering. The moment of departure from a particular cluster was marked by the timestamp of the last GPS sample in a sequence containing only that cluster. Since our target scenario focuses on departures from a user's home, we searched for sequences labeled as `HOME` and identified the last sample in these sequences. This sample's timestamp was determined as the departure time only if a different location label appeared within the next 15 minutes, addressing reliability issues caused by missing data.

The tracking sequence for HOME departures was reset upon the next detection of a HOME label. As a result, 1,747 ground-truth samples were obtained for our departure time analysis, including results from participant exclusion which will be described in Section 5.1.1.

5 Data Preprocessing

This section provides a comprehensive overview of our data preprocessing pipeline, encompassing data cleaning and feature extraction methodologies. We present a robust data exclusion strategy for both participant and passive sensing data to ensure the reliability of our dataset. Leveraging this refined dataset, we establish a feature extraction pipeline that involves engineering passive features, selecting them based on correlation analysis, and implementing a sliding window approach for real-time departure time prediction. The sliding window approach ensures the practicality of the *DFC* algorithm for real-time applications.

5.1 Data Cleaning

Before conducting detailed data analysis, we implemented a two-fold data exclusion strategy to ensure the reliability of our dataset. Mobile passive sensing data collected in the field often suffer from quality issues due to technical problems during data transmission or inattentiveness of end users [1, 31, 61]. To address this, we first excluded participants with low data quality. Additionally, we removed redundant entries to maintain the temporal integrity of passive sensing data and eliminated clearly erroneous data based on realistic ranges. These steps provide a comprehensive framework for refining our dataset, ensuring robust data-driven research.

5.1.1 Participant Exclusion. Data completeness is a crucial metric for data quality, measuring whether all necessary information is available for accurate analysis. Incomplete data can lead to biased results and undermine the reliability and validity of the study. Therefore, we established participant exclusion criteria to ensure data quality. Participants were excluded if they: 1) lacked at least one data entry across all sensors, 2) had data collected for less than 2 months, 3) had fewer than 30 out of 60 samples (days) within this 2-month period, and 4) missed more than one sample for both weekdays and weekends in any given week in the dataset. As a result, a total of 48 participants were retained for our analysis.

5.1.2 Passive Sensing Data Exclusion. Our mobile passive data cleaning method consists of two key strategies. First, we ensure temporal integrity by removing temporarily incorrect or redundant data. Second, we filter each sensor's data to exclude unrealistic values based on real-world ranges. This streamlined data cleaning process guarantees the reliability of our dataset.

- **Temporal integrity preservation:** we ensured the temporal integrity of the data by removing samples that were temporally reversed or duplicated due to rapid sampling. Redundancy can occur in event-based data because of the sensitivity of built-in sensors, user interactions with apps, and issues with data transfer or storage on the server side. These issues affect the reliability of data, such as step counts and unlock/screen states. To address this, we determined a time cut-off by analyzing the histogram of the duration between consecutive samples, identifying a local minimum point that distinguishes expected redundant data from normal data (*e.g.*, 0.1 seconds for step data, 1 second for unlock/screen state).
- **Erroneous data and outlier removal:** we removed clearly erroneous data based on the domain characteristics of corresponding sensor. This involved filtering continuous sensing data such as ambient light (brightness in lux, ranging from 0 to 10,000) and sound (sound energy in dBFS, less than 0) to ensure the values fell within realistic ranges. This data cleansing process ensures the reliable capture of environmental cues related to the departure context by eliminating incorrect outliers.

5.2 Feature Extraction

After cleaning the dataset, we extracted personalized features of individuals to predict departure time in real-time. First, we designed relevant features from continuous sensing and event-based sensors. These features were then selected for each individual based on their statistical significance, as determined by the *p-value* derived from their correlation with ground-truth data. This feature selection process enables the predictive model to learn personalized patterns of passive features related to departure. Finally, we implemented a sliding window technique for dynamic, real-time observation of departure patterns, thereby significantly enhancing the practicality of the *DFC* algorithm.

5.2.1 Feature Engineering. We generated features based on two types of sampling methods: continuous sensing-based data and event-based data. For continuous sensing data, such as ambient light and sound, we calculated various statistical metrics. These include the mean, minimum, maximum, standard deviation, skewness, and kurtosis within a fixed time window to represent the data distribution [40, 55, 63]. For event-based data, which includes behavioral aspects like physical activities, social interactions, and smartphone usage, we extracted both time-domain and frequency-domain features. Time-domain features include the duration in hours for specific types of sensor data (*e.g.*, the duration of walking as detected by activity type), calculated from the time intervals between start and end times or between consecutive samples. Frequency-domain features represent the counts of samples corresponding to each sensor data. These duration and count-based features quantify the amount of behavior each sensor detects, providing insights into individuals' behavioral tendencies prior to departure.

To capture the complex characteristics of different sensors, we developed nonlinear features that reflect their unique properties. For example, entropy measurements in event-based data capture the randomness across different categories [41, 46]. Additionally, we utilized the timestamps denoting the initial occurrences within each sensor data stream, providing crucial insights such as identifying the moment of first activation or engagement related to each factor. For example, the timestamp of the initial screen unlock or interaction is key to determining the start of an individual's daily activities. As a result, we created a comprehensive set of 98 passive features in total. The specifics of these extracted features are systematically detailed in Table 1.

5.2.2 Feature Selection. Feature selection is crucial for enhancing the computational efficiency and predictive performance of our model through personalization. We filtered features for each individual based on statistical significance (*p-value* < 0.05) using Spearman's rank correlation. This coefficient is effective for capturing non-linear yet monotonic relationships between variables that do not follow a normal distribution and is robust to outliers, making it suitable for our dataset. To identify the most informative window duration for capturing departure signals, we evaluated various time window sizes ranging from 1 to 12 hours with a stride of 1 hour (*i.e.*, sliding window), ensuring that each window concluded 30 minutes before the ground-truth departure time (time allocated for fast charging until reaching 100% SOC). Consequently, social activities involving communication app usage, smartphone usage, ambient brightness and sound energy were found to be highly correlated with departure time, with a 6-hour window size being the most effective. More details will be discussed in Section 7.1.

5.2.3 Sliding Window Approach for Real-time Detection. To determine when to resume fast charging for full battery capacity, the *DFC* algorithm continuously monitors behavioral and environmental contexts until the most likely moment for pre-departure detection is identified. As addressed in Section 3.1, it is important to predict this moment 30 minutes prior to departure to ensure sufficient time for fast charging to reach 100% SOC. For real-time departure time prediction, the algorithm leverages a 6-hour sliding window technique to extract mobile passive features at 20-minute intervals.

To enhance real-time departure time prediction, we additionally incorporated current time information to account for regular patterns. This includes the current hour, day of the week, and daily epochs consisting of five representative periods: morning (6 a.m. to 12 p.m.), afternoon (12 p.m. to 6 p.m.), evening (6 p.m. to 9 p.m.),

Table 1. Extracted features from passive sensing data, comprising 98 features in total.

Data	Sampling Rate	Extracted Features	Relevance to Departure Detection
Behavioral data			
Activity Type: still, walking, running, in_vehicle, on_bicycle	Event-based	Durations of 5 activity types Counts of 5 activity types Ratios of 5 activity types to total First sample times of 5 activity types	Physical activities (movements) as indicatives of departure
Step	Event-based	Count of step Std. of step duration First sample time of step	
Significant Motion	Event-based	Count of significant motion Time elapsed since last significant motion First sample time of significant motion	
Call: in, out, missed	Event-based	Durations of in and out call Counts of 3 call types Ratio of in/out call Ratios of 3 call types to total First sample times of 3 call types	Social (call) activities as indicatives of departure
Screen State	Event-based	Durations (mean, total) of screen on state Counts of screen states (on, off) Count of short screen on state First sample time of screen on state Entropy and normalized entropy of screen states (on, off)	Smartphone usage behaviors as indicatives of departure
Unlock State	Event-based	Durations (mean, total) of unlock state Counts of unlock states (unlock, lock) First sample time of unlock state Entropy and normalized entropy of unlock states (unlock, lock)	
App Usage: social, communication, lifestyle, entertainment, health, book, video, finance, photography, productivity	Event-based	Durations of 10 app types Counts of 10 app types First sample times of 10 app types Entropy and normalized entropy of usage across 10 app types	Active app types (chatting, maps, etc) as indicatives of departure
Environmental data			
Ambient Light	1 per 15 minutes	Count of complete darkness, min, max, mean, std., kurtosis, skewness of brightness (lux)	Light intensity and variations as indicatives of departure
Ambient Sound	1 per 15 minutes	Min, max, mean, std., kurtosis, skewness of sound energy (dBFS)	Sound intensity and variations as indicatives of departure

night (9 p.m. to 12 a.m.), and midnight (12 a.m. to 6 a.m.). The days of the week were categorized as weekdays or weekends. These temporal features were then converted into numerical codes, providing contextual timing information based on daily and weekly routines to the model.

6 Outlier Detection for Predicting Departure Time

We develop a robust methodology for continuously predicting EV departure times using a sliding window approach. Departure occurrences are treated as anomalies, given their specific timing within a day. To effectively detect these outliers, we adopt Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [14], further enhanced with kneedle approach [49]. Our DBSCAN-based outlier detection method enables precise, real-time prediction of departure times, facilitating personalized and adaptable EV charging strategies that align with individual lifestyles.

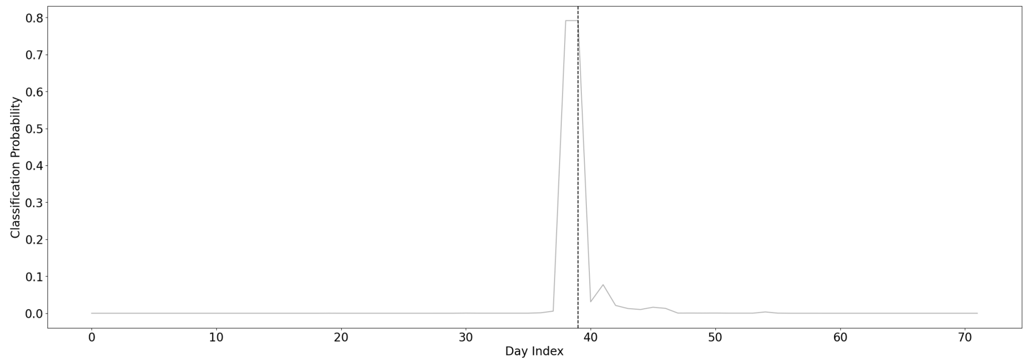
6.1 Departure as an Outlier

In this study, the target scenario for the *DFC* algorithm involves charging EVs overnight at home and departing the next day. Therefore, we focus on detecting the first significant departure of the day. In other words, all times except for the first departure are considered non-departures. The sliding window method significantly increases the data sample size, with each successive 6-hour window moving in 20-minute strides and overlapping approximately 95% of the previous window. This approach generates 72 samples over a 24-hour period, resulting in a significant label imbalance. Additionally, since we only consider days when departures occur, at least one departure must be detected each day. To address these issues, we first examine the classification probability of each sample within a day to determine the optimal departure classification threshold. We then present DBSCAN-based method to detect this threshold as an outlier, thereby identifying the moment of departure.

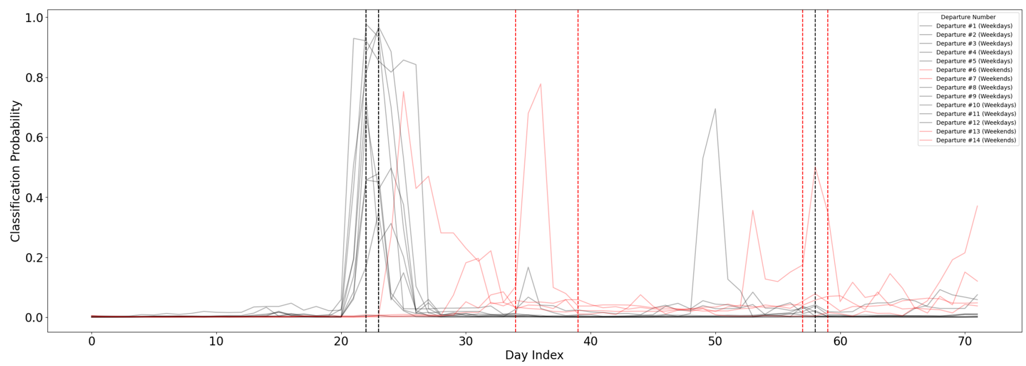
To obtain the classification probability distributions, we evaluated the tree-based ensemble models based on Leave-One-Day-Out Cross Validation (LODO CV) with a 14-day training set. Detailed descriptions of our evaluation settings are provided in Section 7.2. Ideally, a well-trained model would show a single peak at the actual departure time in the classification probability distribution throughout the day. This peak would imply the model's recognition of behavioral, environmental, temporal indicators of departure. Figure 5(a) displays an example of a classification probability distribution for a particular weekday. Two distinct probability peaks, along with the dotted line indicating the ground-truth reference around 1 p.m. suggest the individual's regular weekday routine. The predictive models can easily detect this moment as departure.

However, practical challenges arise due to the irregularity of departure times, particularly on weekends. Departures may occur multiple times in a day, resulting in multiple peaks within the daily probability distribution, creating ambiguity about which peak represents the right departure. Figure 5(b) illustrates the distribution of departure times across all 14 days, with weekdays and weekends distinguished by color. Weekday distributions, marked in grey are relatively concentrated around 7 a.m., indicating consistency in departure times. In contrast, weekend distributions depicted in red show several peaks at various times throughout the day. The classification distribution for a given day is unknown until the *DFC* algorithm processes and produces the probabilities for all time points in that day. Therefore, to avoid reducing the driving range of EVs due to late predictions (*i.e.*, failed operation as shown in Figure 2(b)) while minimizing $t_{100\%}$, we consider the first peak as the moment of departure. This approach balances the minimizing battery degradation with maintaining user acceptance of the *DFC* algorithm.

Furthermore, using a standard classification probability threshold (*i.e.*, 0.5) to identify departure events may not be feasible, as it does not assure the detection of at least one departure per day. As shown in Figure 5(b), several peak probabilities may not reach the 0.5 threshold, leading to missed departures on certain days. The goal is to pinpoint the moment of sharp rise and subsequent fall in probability around the time of departure among the 72 samples. Regardless of the magnitude of the classification probability, this moment of sudden increase can be captured using an outlier detection method.



(a) Classification probability distribution of a single weekday



(b) Classification probability distributions of weekdays and weekends

Fig. 5. Classification probability distribution of a single weekday.

6.2 DBSCAN-based Outlier Detection

Our approach to predict departure times uses a DBSCAN-based outlier detection method, inspired by control chart methodologies commonly used in industrial engineering for quality control. Control charts (also known as Shewhart charts) [54] are widely used in industrial engineering due to their effectiveness in detecting sudden shifts in time series data in real-time, which aligns with our objective of identifying departure events. However, these charts require adaptation to the variable nature of daily life, necessitating a self-adjusting mechanism for thresholds or control limits. We initially implemented a simplified version of our algorithm, where the control limits are dynamically adjusted using historical data. This setup provides the foundation for a more refined, self-adaptive strategy that employs DBSCAN to calculate control limits.

DBSCAN's strength in detecting outliers in time series data lies in its density-based clustering mechanism, which evaluates the local neighborhood density of data points. This is defined by parameters epsilon (ϵ) and minimum points (*MinPts*), used to identify core samples and outliers [14]. This method is well-suited for time series data due to its flexibility in forming clusters of varied shapes without predefined cluster sizes and its ability to identify significantly deviating probability samples using squared Euclidean distance. Given the daily departure variations both within individuals (intra-personal) and across different individuals (inter-personal), a static configuration for DBSCAN's parameters is insufficient. Instead, we incorporate the kneedle algorithm [49] for automatic tuning of ϵ and *MinPts*, ensuring DBSCAN's parameters are optimally set to reflect each individual's unique departure patterns. This method involves flipping the k-dist graph to transform its concavity, rotating it until the line formed by the endpoints aligns with the x-axis, and identifying the peak of the curve as the knee point. Our adaptive parameter tuning enables personalized and dynamic adjustment of DBSCAN parameters, enhancing outlier detection accuracy amidst the complexities of daily departures [47].

Upon implementing this methodology on a training set, we determined an optimal *MinPts* that effectively minimizes the temporal discrepancy between the occurrence with the highest probability and the initial outlier identified by the DBSCAN algorithm. This *MinPts* is customized for each individual to accommodate inter-personal variability. The kneedle algorithm enables adaptive adjustment of *MinPts* for each departure instance in subsequent, unobserved test datasets, thereby addressing intra-personal variations. Predicting departure times during weekends presents additional complexities due to irregular activities such as trips or social gatherings, which affect classification probability distributions. This differential analysis ensures that the *DFC* algorithm significantly enhances its accuracy and adaptability across different day types. Overall, this approach represents a substantial advancement in departure time prediction, effectively handling both regular and irregular, yet predictable behavioral patterns.

Algorithm 2 outlines a comprehensive method for predicting departures through outlier detection. The process begins with a pre-trained machine learning model (tree-based ensembles) generating classification probabilities as it traverses the dataset using a sliding window. A crucial step involves accumulating a sufficient number of probability samples (exceeding K, an individual-specific minimum threshold set at 12 to account for typical departures occurring after 4 a.m.) for robust outlier detection. A probability exceeding the conventional 0.5 threshold indicates a departure event. If this threshold is not met, the kneedle algorithm is applied to a k-distance graph customized with a personalized K to determine the optimal ϵ. DBSCAN is then used to identify a potential outlier which is considered as the departure moment, by gradually increasing K until an outlier is detected. This approach provides our algorithm with the adaptability needed to effectively handle new, unseen datasets.

7 Evaluation

Our evaluation framework is structured into three distinct phases. First, we analyze features correlated with departure time using Spearman's rank correlation to select relevant features (Section 7.1). Second, we evaluate our digital phenotyping approach against a historical-based approach for comparative analysis. We demonstrate that our method outperforms the baseline by capturing irregular patterns indicative of behavioral and environmental contexts (Section 7.2). Finally, we implement adaptive departure time prediction using an online learning framework to enhance the practicality of the *DFC* algorithm for real-world deployment (Section 7.3). We split the dataset into a 14-day training set for correlation analysis and comparative evaluation, followed by a subsequent 14-day test set for performance evaluation of adaptive online learning. Ultimately, our findings underscore the validity and feasibility of utilizing a digital phenotyping approach to enhance departure time prediction within the *DFC* algorithm.

7.1 Correlation Analysis

Feature selection process described in Section 5.2.2 aims to filter personalized features correlated with an individual's departure time using Spearman's rank coefficient. The correlation analysis reveals that passive features are capable of capturing departure time and identify behavioral and environmental categories with a general tendency for high correlation.

Figure 6 displays a heatmap of statistical significance (left) and boxplots of correlation coefficients (right) for the extracted features, which are sorted in descending order based on the highest average coefficients across all participants. Various time window sizes, ranging from 1 to 12 hours with a 1-hour stride, are used for weekday

Algorithm 2 Departure Detection Algorithm

```
Require: x_test, model, samples_per_day, MinPts, index (index > MinPts)
Ensure: pred_time
prob_list ← []
pred_time ← 0
for i ← 0 to index − 1 do
    prob_list.append(model.predict(x_test[i]))
end for
for i ← index to samples_per_day do
    MinPts_day, outlier, ϵ ← MinPts, 0, 0
    prob ← model.predict(x_test[i])
    prob_list.append(prob)
    if prob ≥ 0.5 then
        pred_time ← x_test[i].time + 1
        break
    end if
    while outlier = 0 do
        k_dist, max_curv ← calc_k_dist(prob_list, MinPts_day)
        knee ← find_max(max_curv)
        ϵ ← k_dist[knee]
        if (MinPts_day = len(prob_list) - 1) or (ϵ = 0) then
            break
        end if
        if ϵ ≠ 0 then
            clust ← DBSCAN(prob_list, ϵ, MinPts_day)
            outlier ← clust.label_last
            if outlier = −1 then
                pred_time ← x_test[i].time + 1
                break
            end if
        end if
        MinPts_day ← MinPts_day + 1
    end while
end for
return pred_time
```

departures. Similarly, Figure 7 displays the same visualizations for weekend departures. The heatmap illustrates the number of participants whose features achieve statistical significance (*p-value* < 0.05) for each time window. A distinctive gradient distribution is observed, with a lighter hue at the center top indicating higher counts, and progressively darkening as the counts decrease away from the center. The boxplots show the range of correlation coefficients for each feature across different time windows, highlighting the predominance of the grey box (*i.e.*, 6-hour window) on the far right for each feature. These results suggest that a medium-length time window of 6 hours is the most effective duration for capturing overall pre-departure contexts through passive data.

We evaluate the similarities and differences between weekday and weekend departure correlation results. Figures 6 and 7 show that the top 10 most correlated features include all five categories of behavioral and environmental

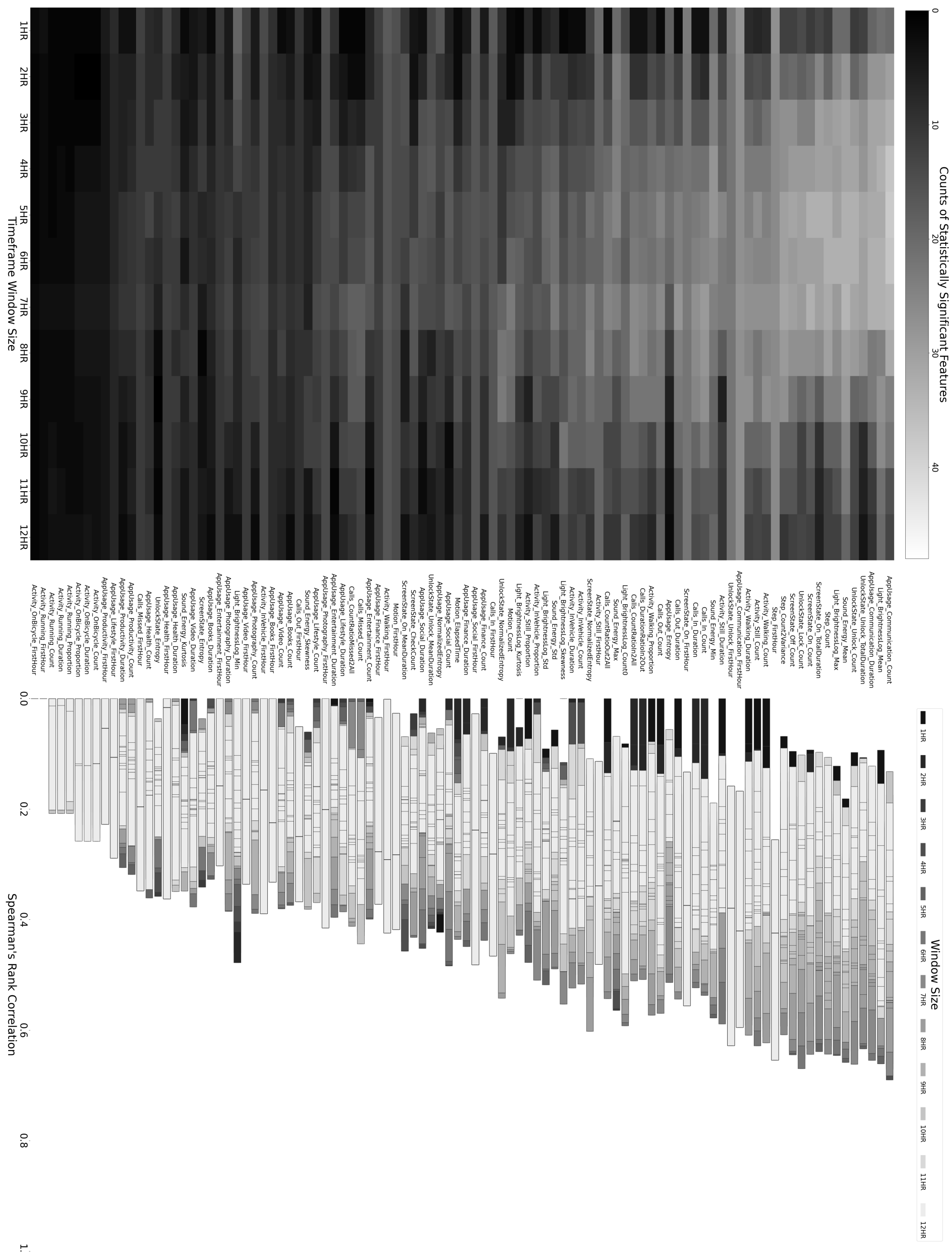


Fig. 6. Spearman's rank correlation coefficient results for weekday departures. The figure on the left presents a heatmap indicating the number of participants with statistically significant features (*p-value* < 0.05), while the figure on the right displays boxplots of correlation coefficients, highlighting the interquartile range for features sorted by average coefficients across participants. Notably, a 6-hour time window effectively captures pre-departure contexts through passive data.

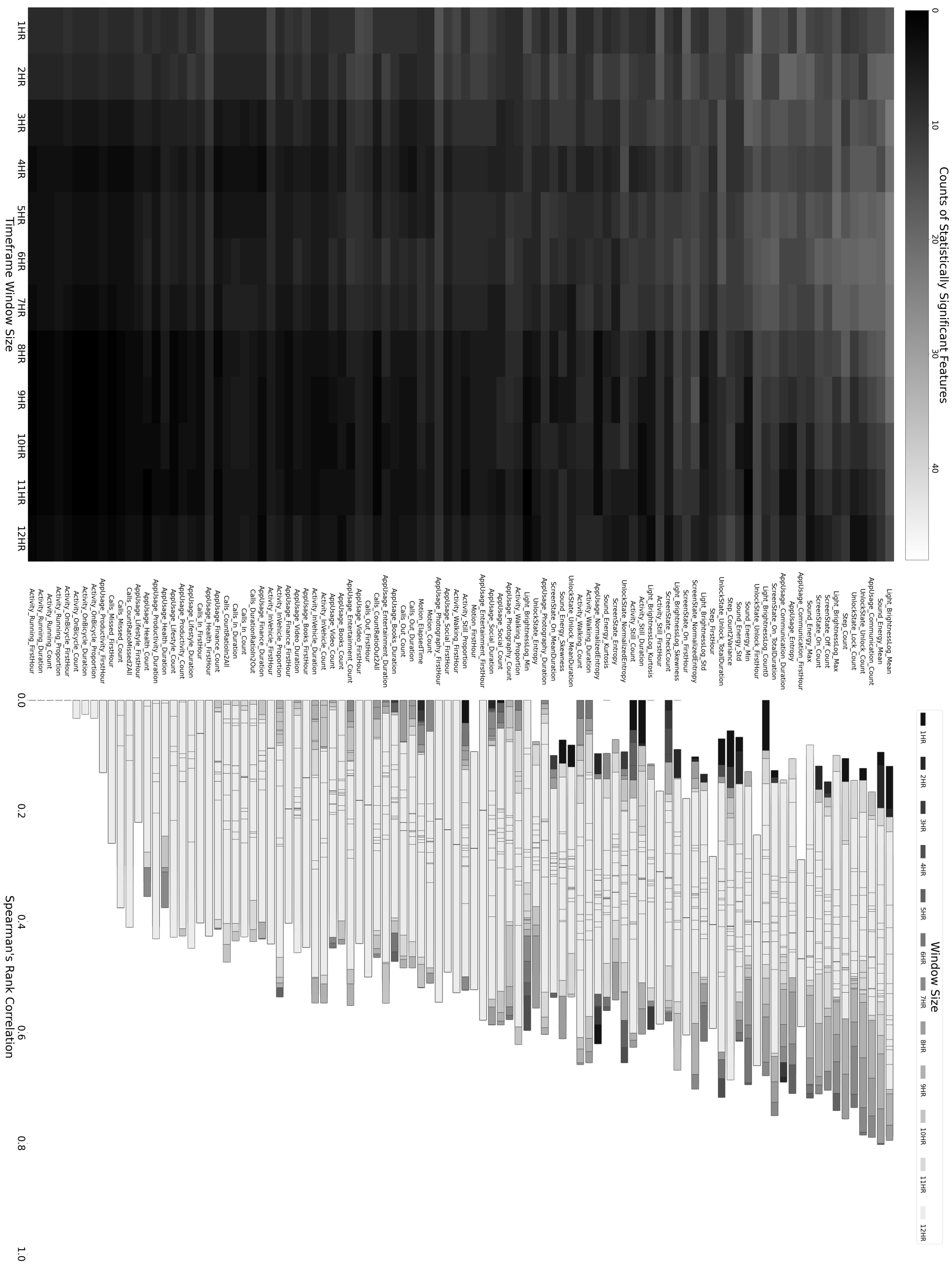


Fig. 7. Spearman's rank correlation coefficient results for weekend departures. Compared to weekdays, the coefficients are stronger, but the statistical significance is relatively weaker. This suggests that weekend departures tend to be more personalized, rather than following the generalized patterns typical of regular weekday routines.

indicators we defined: physical activities, social interactions, smartphone usage, ambient light, and ambient sound. Notably, `Step_Count`, `AppUsage_Communication_Count`, `UnlockState_Unlock_Count`, `Light_BrightnessLog_Mean`, and `Sound_Energy_Mean` consistently rank among the top features for both weekdays and weekends, indicating their representativeness across all categories. The weekday correlation results show higher statistical significance, as indicated by the brighter hues in the heatmap, but exhibit relatively lower coefficients in the box-plots. In contrast, the weekend results display the opposite pattern. These findings suggest that weekday departure patterns reflect more generalized tendencies linked to individuals' contexts due to their regular characteristics, while weekend departures exhibit a stronger relationship but in a more personalized manner.

7.2 Comparative Evaluation of Departure Time Prediction

We evaluate our digital phenotyping approach against a baseline model that relies on temporal patterns by analyzing the histogram of predicted departure times in the training set. Our analysis demonstrates that mobile passive features indicative of behavioral and environmental contexts more effectively capture the variability of departure patterns compared to traditional historical-based methods.

To achieve this, we use tree-based ensemble models (*i.e.*, XGBoost, CatBoost, LightGBM, and RandomForest) to train on mobile passive features, as presented in Table 1. These models handle diverse data types (*e.g.*, numerical, categorical) and mitigate overfitting by leveraging multiple trees. For baseline comparison, we adopt LSTM network, which is effective in time series forecasting by utilizing historical data. LSTM is well-suited for capturing long-term dependencies within sequential data, making it an appropriate benchmark for highlighting the temporal intricacies involved in departure time prediction. Our LSTM network consists of a single LSTM layer with 32 units, followed by a dense layer with 2 units and a softmax activation for binary classification. The model is trained for 100 epochs with a batch size of 72, using categorical cross-entropy loss and the Adam optimizer. To contrast the efficacy of digital phenotyping in capturing irregular departure patterns, the LSTM model solely learns temporal features described in Section 5.2.3.

For evaluation, we use LODO CV, excluding one day at a time to predict departure times. The LODO CV results on the training set are illustrated using Kernel Density Estimation (KDE) plots across the various models. The KDE plot estimates the probability density function of data with a smoothing kernel, producing a continuous curve that reveals its underlying structure. Figure 9 displays the KDE outcomes for tree-based models, which incorporate behavioral, environmental, and temporal features, alongside the LSTM model that utilizes historical features, with ground truth as a reference. We compare our proposed model with the baseline by explicitly showing how each model aligns with the actual ground-truth distribution for weekday and weekend scenarios. The x-axis denotes the time of day transformed into a day index derived from a sliding window consisting of 72 resolutions, while the y-axis indicates the probability density of departures occurring at each specific day index.

Upon examining the ground-truth data in Figure 9, we observe a distinctive spike around the day index corresponding to the morning period on weekdays, suggesting regular commuting departure patterns. In contrast, the weekend distributions show a more normal distribution, indicating that departures are more irregular and less predictable due to greater variability. The KDE results of the tree ensemble models trained with mobile passive features more closely resemble the ground truth compared to those of the LSTM model. This suggests that there are identifiable patterns through behavioral and environmental indicators which were previously considered irregular when viewed solely from a temporal perspective, especially during weekends. The LSTM model, however, struggles to clearly distinguish between weekdays and weekends in its probability density distributions. This lack of distinction is likely due to the dataset's bias towards weekday regularity, which significantly influences the LSTM's learning from historical data. Overall, these results demonstrate the effectiveness of leveraging behavioral and environmental indicators along with temporal features for more accurate departure time predictions.

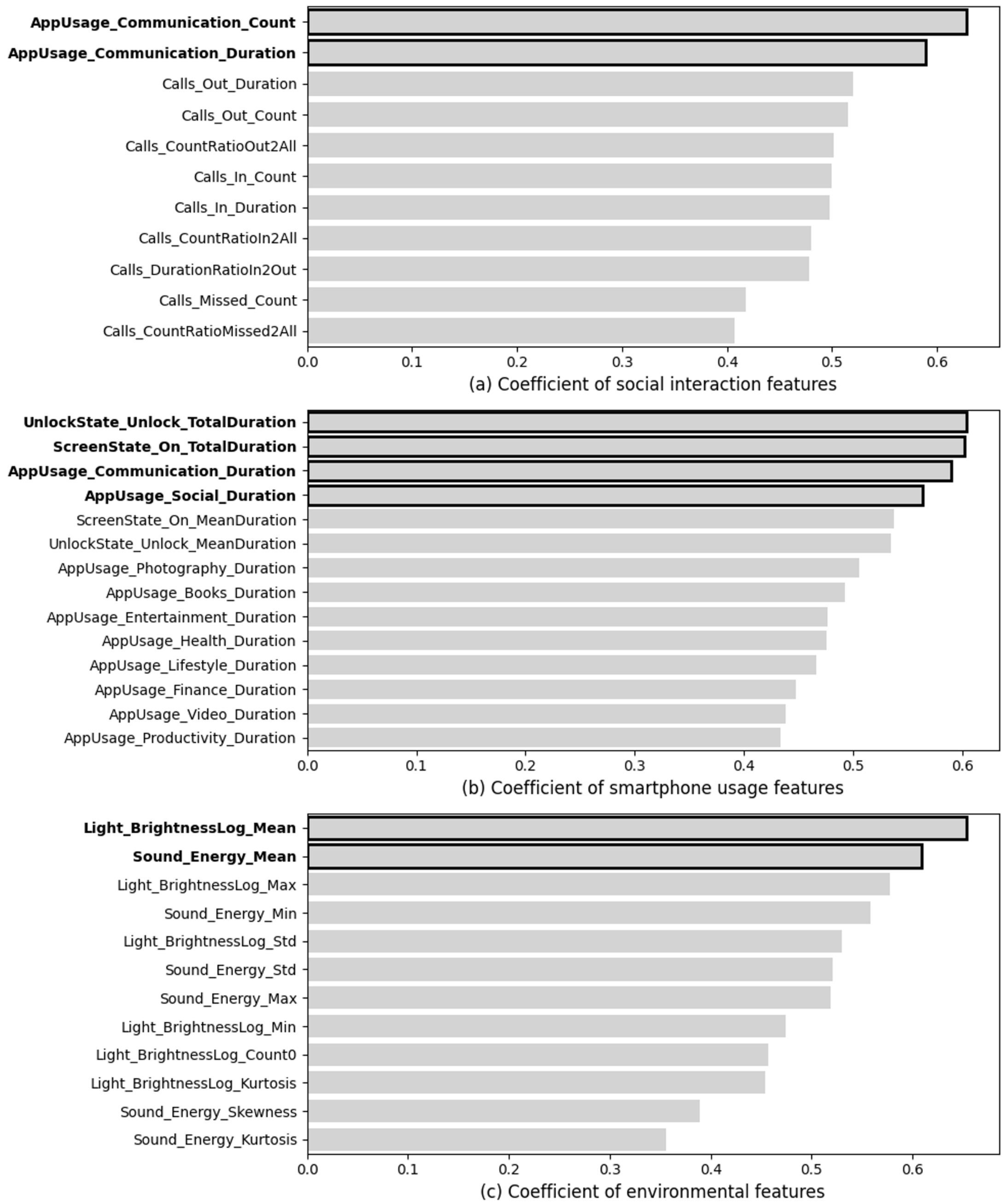


Fig. 8. Bar graphs depicting correlation coefficients for features related to (a) social interactions, (b) smartphone usage, and (c) environmental contexts. Communication app usage has a higher correlation than call-related features, indicating a preference for chatting over calls before departure. Smartphone usage is primarily correlated with communication activities. Simple representations of ambient light and sound (mean values) effectively capture the environmental context preceding departure. Significant features within each category are marked with highlighted outlines on the bar graphs.

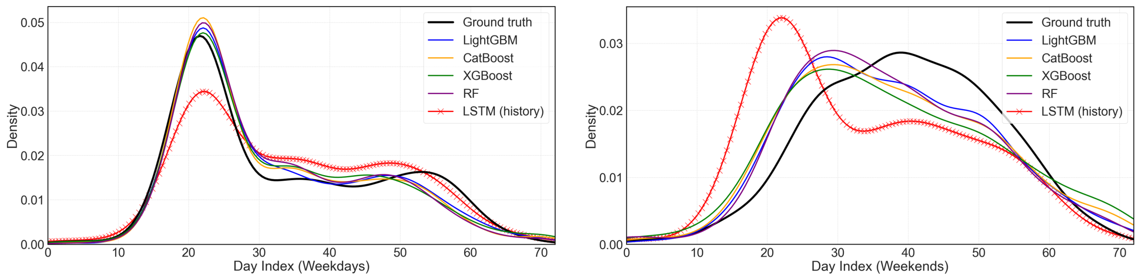


Fig. 9. Kernel Density Estimation (KDE) plot of probability over day index across different models. Compared to the weekdays (upper), the weekends (lower) display a normal distribution, underscoring the presence of irregular patterns.

7.3 Adaptive Online Learning Using DBSCAN

Lastly, we validate our proposed method on a test set to verify its effectiveness in predicting departure times. To develop a robust approach that handles dynamic and adaptive departure patterns, we incorporate an online learning technique, which enhances the model's performance as more data becomes available. We adopt MAE measured in hours as the performance metric to quantify the direct deviation between predicted and actual departure times, providing a clear measure of model error. Given the importance of minimizing undesired predictions to ensure EV driving range while maintaining departure time prediction accuracy, we introduce a Margin Estimation Plot (MEP). This tool assesses the necessary time margin to effectively eliminate the failure cases caused by late predictions, further enhancing our model's robustness.

Our adaptive online learning approach for real-time departure time prediction follows the process outlined in Algorithm 2. Initially, tree-based ensemble models are pre-trained using a 14-day training set, followed by predicting departure time in one-day increments over a subsequent 14-day test set. As the sliding window progresses, the tree-based ensemble models generate classification probabilities for each input instance. These aggregated probabilities are then fed into DBSCAN to identify the potential departure moment by detecting probability outliers. At this stage, *MinPts* is optimized for each individual to minimize the MAE, resulting in more personalized departure time prediction. This method allows us to observe how the *DFC* algorithm adaptively manages probability distributions and evolves as more data is accumulated over time.

Table 2 presents the MAE results of departure time predictions for weekdays, weekends, and all days across four ensemble models. Overall, LightGBM performs best, achieving an MAE of 2.18 on weekdays and 4.46 on weekends for departure detection. Error rates during weekends are 50% to 100% higher than on weekdays, due to the irregular patterns of individual departure behaviors. This explicitly shows the challenge of predicting weekend departures, as they contain inherent randomness. Moreover, the models did not show significant improvement through incremental learning, which continuously integrates new data for training. The task of predicting departures as outliers is predominantly influenced by DBSCAN rather than the machine learning model itself, due to the model's inability to effectively learn the probability distribution.

As explained in Section 3.1, we use t_{margin} to determine the extra time required for fast charging when the predicted departure time is later than the actual departure time, ensuring the maximum driving range. To effectively address this, we design a Margin Estimation Plot (MEP) as a visualization tool to assess the accuracy and uncertainty of the predictive model. The MEP graphically displays the difference between predicted and true departure times using scatter plots with a reference line. This baseline intuitively shows the magnitude and direction of by which each prediction deviates from the true departure time. Figure 10 presents four examples of

Table 2. Performance results across tree-based ensemble models (MAE).

	LightGBM	CatBoost	XGBoost	RandomForest
All days	2.76	3.55	3.13	4.13
Weekdays	2.18	3.06	2.57	3.62
Weekends	4.46	4.97	4.78	5.62

MEPs using LightGBM results, illustrating how t_{margin} is determined for individuals and the overall distribution of prediction deviations from the ground truth.

Figure 10(a), (b) and (c) show three representative cases of MEP for a particular participant. The X and Y axes represent the predicted and true departure times in hours respectively, with a solid line (y=x) indicating the ideal reference. Each annulus point represents a daily departure sample in the test set, color-coded to distinguish between desired predictions (predicted earlier, shown in yellow) and undesired predictions (predicted later, shown in purple) compared to the actual time. Points marked in purple to the right of the line indicate undesired predictions that are later than the actual time. By applying the t_{margin} to the reference line, these undesired predictions can be relatively shifted to the left of the new dotted reference line. In Figure 10(a) (PID 583), all departure times are predicted before the true time, successfully preventing undesired predictions. In this case, t_{margin} is not needed to ensure sufficient charging time. On the other hand, Figure 10(b) (PID 599) and Figure 10(c) (PID 185) depict scenarios where t_{margin} needs to be applied. The time difference in hours between the solid and dotted lines indicates the t_{margin}, which is small for PID 599 (0.34 hours) and large for PID 185 (2.34 hours). In addition estimating t_{margin}, MEPs demonstrate the effectiveness of our method in learning irregular departure patterns by capturing behavioral and environmental contexts via digital phenotyping. For example, in the case of PID 185, while most points cluster around 07:30, the point near 13:00 is close to the solid reference line, and the overall trend of all annulus points aligns with the line.

Figure 10(d) extends this analysis to encompass all participants, providing prediction results for both weekdays and weekends, differentiated by color coding (black for weekdays, red for weekends). While weekday predictions tend to be more accurate than weekend predictions, generally aligning more closely with the solid reference line, there is still some variability. This variability is caused by irregular weekday scenarios such as holidays or occasional trips. Similar to the MEPs for individual participants, points on the left side indicate predictions that are earlier than the true departure time, whereas points on the right side indicate later predictions. While reducing variability on both sides is important for minimizing overall model uncertainty, our primary goal is to prevent insufficient charging due to late predictions on the right side. By applying the optimized t_{margin} for each user based on prediction uncertainty, we can ensure sufficient driving range and enhance user acceptance. Further discussion on this will be covered in Section 8.3.

8 Discussion

We first discuss the responsibilities of individuals and manufacturers regarding the successful implementation of the *DFC* algorithm and privacy protection (Section 8.1). Next, we elaborate on the privacy measures and the ethics processes involved in our study (Section 8.2). We also discuss the practicality and user acceptance, highlighting the solutions we implemented (Section 8.3). Lastly, we discuss the limitations of our work and outline future research directions (Section 8.4).

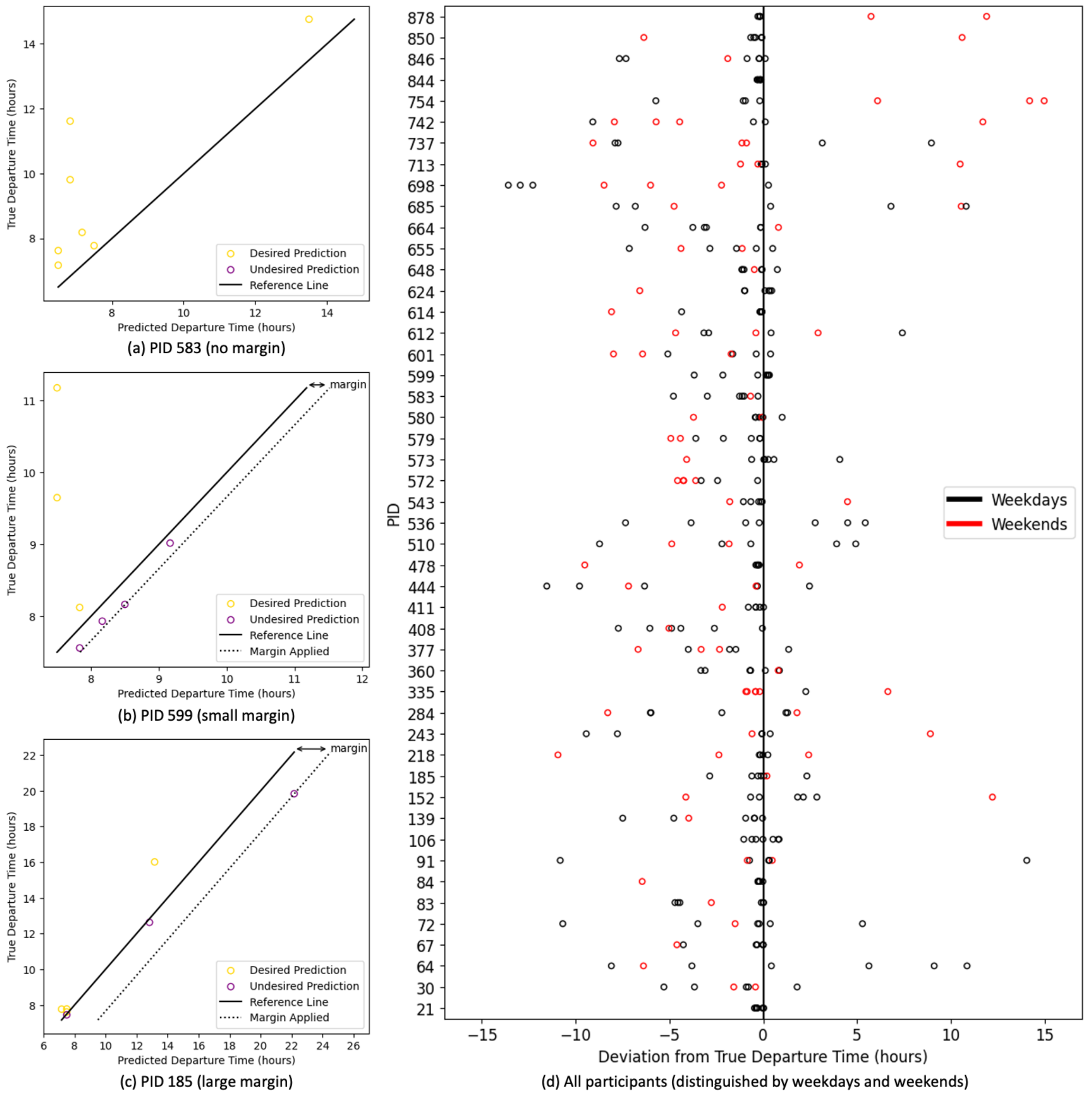


Fig. 10. Margin Estimation Plots (MEPs) of four representative scenarios. MEPs display the differences between predicted and true departure times using scatter plots with a reference line. (a), (b), and (c) show the examples for a particular participant, while (d) presents the results for all participants. Based on the MEPs, we adopt an optimal t_{margin} as a safeguard to ensure sufficient EV driving range.

8.1 Responsibilities of Individuals and Manufacturers

The successful implementation of the *DFC* algorithm involves various stakeholders in the EV industry, namely EV users and manufacturers. The *DFC* algorithm utilizes lifelogging data collected from smartphones in real-time, enabling automatic EV charging without disrupting users. The performance of the *DFC* algorithm can be enhanced through adaptive online learning, as the predictive model is personalized based on accumulated departure data samples. Consequently, the *DFC* algorithm's effectiveness heavily depends on the quality of individual data. Thus, EV users are responsible for data quality by tracking their data in daily lives, which includes ensuring the data collection app, EVA, remains active and reporting any issues with data collection to the service provider. Users can monitor and address data collection issues using data quality management features in EVA, such as calendar interface and data-uploading function. On the other hand, EV manufacturers can integrate the *DFC* algorithm into BMS software of their EV products. Their objective is to provide a robust service to users and thereby expand their market. However, since they handle personal data of EV users, they are responsible for ensuring the privacy, security, and ethical use of the data. This includes obtaining explicit user consent, complying with regulations like General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA), encrypting data, implementing access controls, and conducting regular security audits [43, 57]. They must use data solely for stated purposes, maintain transparency, and allow users to access, correct, and delete their data.

8.2 Privacy and Ethics

This study was approved by our IRB and adhered to established ethical guidelines. Participants provided informed consent after being fully briefed on the study's objectives, methods, and potential risks and benefits, and were compensated up to $50 based on their participation. Data were collected using a custom mobile application, EVA, with all information encrypted using AES-256-CBC and securely stored on a Linux-based server to ensure privacy through anonymization. Data and personal information are retained for up to five years for follow-up research complying with the Personal Information Protection Act and are securely disposed of afterward. Strict data usage and access controls are enforced to prohibit data sharing outside the research team and prevent commercial use, although anonymized sample data may be provided to thesis judges for transparency. Institutional committees and government authorities have access to review personal information within regulatory boundaries to ensure data reliability and integrity. Our study implemented rigorous privacy protection measures to safeguard participant data and ensure compliance with ethical standards and legal requirements, thereby guaranteeing data confidentiality and integrity throughout the research process.

8.3 Practicality and User Acceptance

Since the *DFC* algorithm leverages smartphone data from individuals, several considerations must be addressed for its practical deployment and user acceptance. The departure time prediction model operates in real-time, starting when the SOC reaches 80% and continuing until a departure is detected within the day. This requires continuous collection and processing of multi-dimensional data from smartphones, leading to increased battery consumption and posing a potential practicality challenge for the *DFC* algorithm. To resolve this, the EVA app minimizes interference by stopping data collection when the phone battery drops below 20% and by running in the background. For real-world application, computationally intensive tasks like departure time prediction will be offloaded to the cloud BMS server, enabling real-time processing of large-scale data. The next-generation cloud BMS can remotely control EV charging by simultaneously monitoring human states from smartphones and EV state in real-time. Additionally, critical sensors relevant to departure time (*e.g.*, app usage, unlock state, ambient light, ambient sound, steps) can be identified based on correlation analysis results (Section 7.1), allowing the exclusion of less important sensors to reduce computation and conserve battery.

Along with the battery consumption issues discussed above, data privacy is a significant factor affecting user acceptance. Ensuring user trust in privacy protections is crucial for adoption of the *DFC* algorithm. We have complied with privacy measures outlined in Section 4.1 and further detailed in Section 8.2. Additionally, inaccurate departure time predictions, especially those predicting later departures than actual time can limit trip distances by not fully charging the EV, leading to user dissatisfaction. To mitigate this, we adopted a time margin (t_{margin}) to account for prediction uncertainty, ensuring sufficient charging time before the actual departure. On the other hand, some users may be reluctant to adopt t_{margin} due to concerns that the potentially extendable $t_{100\%}$ after fast charging may diminish the effectiveness of battery degradation mitigation, even if it results in a less full charge. Therefore, conducting a user survey on the necessity of t_{margin} can offer valuable insights into user preferences between prioritizing battery longevity and ensuring a fully charged battery, which remains one of our future work.

Our adaptive online learning method continuously refines the departure time prediction model by accumulating data and training personalized model for individuals. By collecting sufficient environmental sensing data (light and sound) throughout the year, the *DFC* algorithm can effectively adapt to seasonal variations, thereby enhancing its practicality and accuracy. Moreover, we will develop a comprehensive strategy based on the Extended Technology Acceptance Model (TAM2) [58] by analyzing factors that contribute to user acceptance beyond the original constructs of perceived usefulness and perceived ease of use. This approach will promote initial acceptance and ensures the long-term integration and deployment of the *DFC* algorithm, fostering sustained user adoption.

8.4 Limitations and Future Work

As an initial attempt to mitigate EV battery degradation through digital phenotyping, the *DFC* algorithm has several limitations. First, we assume that the departure patterns of non-EV users do not substantially deviate from those of EV users. However, focusing specifically on EV users would likely yield more precise results. Additionally, collecting more demographic information about participants, such as income, occupation, marital status, and geographic location, could provide deeper insights and more comprehensive findings regarding departure times. From a data perspective, we relied on GPS data to determine ground-truth departure times, which may not accurately reflect actual departure times. This is because the location clusters assigned to each GPS sample and the corresponding marked times are highly dependent on the clustering method (POI) and its sampling rate. Lastly, we used standard models (tree-based ensembles) tp predict the classification probability of departures and DBSCAN to detect probability outliers, rather than designing a dedicated model optimized for learning departure patterns from passive sensing. The current method does not yet guarantee reliable detection of departures, and significant improvements in departure time prediction performance are required to ensure practical applicability in real-world scenarios.

To address these issues, our future work will involve a user study with approximately 500 EV users and establishing a large-scale dataset that integrates human data and EV data. Human data will include diverse demographic information and passive sensing data, while EV data will consist of BMS data (*e.g.*, SOC, voltage, current, temperature) and sensor data (*e.g.*, GPS, accelerometer, gyroscope, atmospheric pressure). For reliable ground truth collection, we will utilize Bluetooth and Android Auto communication between the smartphone and the vehicle to obtain connection times, accurately capturing the exact moment of EV departure. By doing so, we aim to enhance the validity of our *DFC* algorithm for departure time prediction. Furthermore, we will develop an integrated reinforcement learning (RL)-based *DFC* algorithm to control decisions on pausing and resuming charging, surpassing the current fragmented use of standard methods for departure time prediction and delayed charging mechanism.

9 Conclusion

In this paper, we propose the *DFC* algorithm to mitigate battery degradation by delaying the full charging of EVs based on predicted departure time (t_{unplug}). Our departure time prediction method leverages smartphone passive features that capture behavioral, environmental contexts, further enhanced with DBSCAN and online learning to improve robustness and adaptability. We validated our departure time prediction method using a real-world dataset of 48 subjects over 28 days. We identified the top 10 most correlated features across five categories, highlighting the prominence of communication app usage and the impact of simple environmental features before departure. Our method outperformed the LSTM baseline by using tree-based ensemble models trained on passive smartphone data, effectively capturing irregular departure patterns. The adaptive online learning approach combined with DBSCAN achieved an average MAE (in hours) of 2.18 for weekdays and 4.46 for weekends. The MEPs derived from these results provide personalized t_{margin} for each user, ensuring sufficient charging time to preserve driving range. To further improve the dataset's reliability and the *DFC* algorithm's robustness, we plan to conduct a large-scale user study with approximately 500 EV drivers over a period of two months and design a dedicated RL-based *DFC* algorithm as part of future research. We anticipate that this study will deepen our understanding of irregular departure patterns and related factors, significantly enhancing departure time prediction performance and supporting the successful implementation of the *DFC* algorithm.

Acknowledgments

This work was supported by Samsung Research Funding & Incubation Center of Samsung Electronics under Project Number SRFC-MA2202-03.

References

[1] Saeed Abdullah, Elizabeth L Murnane, Mark Matthews, Matthew Kay, Julie A Kientz, Geri Gay, and Tanzeem Choudhury. 2016. Cognitive rhythms: Unobtrusive and continuous sensing of alertness using a mobile phone. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 178–189.

[2] Apple Inc. 2023. About Optimized Battery Charging on your iPhone. https://support.apple.com/en-us/108055. Accessed: 2023-07-06.

[3] Peter M. Attia and et al. 2022. "Knees" in lithium-Ion battery aging trajectories. *Journal of The Electrochemical Society* 169, 6 (2022).

[4] Perla B. Balbuena and Yi Xuan Wang (Eds.). 2004. *Lithium-ion batteries: solid-electrolyte interphase*. World Scientific.

[5] Alexandre Beaudet, François Larouche, Kamyab Amouzegar, Patrick Bouchard, and Karim Zaghib. 2020. Key Challenges and Opportunities for Recycling Electric Vehicle Battery Materials. *Sustainability* 12, 14 (July 2020), 5837. https://doi.org/10.3390/su12145837

[6] Mouaad Boulakhbar, Markos Farag, Kawtar Benabdelaziz, Malika Zazi, Mohamed Maaroufi, and Tarik Kousksou. 2022. Electric Vehicles Arrival and Departure Time Prediction Based on Deep Learning: The Case of Morocco. In *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*. 1–8. https://doi.org/10.1109/IRASET52964.2022.9738115

[7] Irmgard Buchberger et al. 2015. Aging analysis of graphite/LiNi1/3Mn1/3Co1/3O2 cells using XRD, PGAA, and AC impedance. *Journal of The Electrochemical Society* 162, 14 (2015), A2737.

[8] Andrew Bwambale, Charisma F. Choudhury, and Stephane Hess. 2019. Modeling departure time choice using mobile phone data. *Transportation Research Part A: Policy and Practice* 130 (2019), 424–439.

[9] Noelia Caceres et al. 2012. Traffic flow estimation models using cellular phone data. *IEEE Transactions on Intelligent Transportation Systems* 13, 3 (2012), 1430–1441.

[10] Chien-Hsin Chung et al. 2020. Optimization of electric vehicle charging for battery maintenance and degradation management. *IEEE Transactions on Transportation Electrification* 6, 3 (2020), 958–969.

[11] Nathan Eagle and Alex (Sandy) Pentland. 2006. Reality Mining: Sensing Complex Social Systems. *Personal and Ubiquitous Computing* 10, 4 (2006), 255–268.

[12] Economic Commission for Europe Inland Transport Committee World Forum for Harmonization of Vehicle Regulations. 2014. Proposal for a global technical regulation on the Worldwide harmonized Light vehicles Test Procedure (WLTP). 162nd session, Geneva, 11-14 March 2014. Submitted by the Working Party on Pollution and Energy.

[13] Akram Eddahech, Olivier Briat, and Jean-Michel Vinassa. 2015. Performance comparison of four lithium–ion battery technologies under calendar aging. *Energy* 84 (2015), 542–550.

[14] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. 226–231.

[15] Bloomberg New Energy Finance. 2022. Electric Vehicle Outlook 2022.

[16] Zohreh Fotouhi et al. 2019. A general model for EV drivers' charging behavior. *IEEE Transactions on Vehicular Technology* 68, 8 (2019), 7368–7382.

[17] Oliver Frendo, Nadine Gaertner, and Heiner Stuckenschmidt. 2021. Improving Smart Charging Prioritization by Predicting Electric Vehicle Departure Time. *IEEE Transactions on Intelligent Transportation Systems* 22, 10 (2021), 6646–6653. https://doi.org/10.1109/TITS.2020.2988648

[18] Chen Gao et al. 2019. Privacy-preserving cross-domain location recommendation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–21.

[19] Jie Han, Arash Khalatbarisoltani, Yalian Yang, and Xiaosong Hu. 2023. Energy Management in Plug-In Hybrid Electric Vehicles: Preheating the Battery Packs in Low-Temperature Driving Scenarios. *IEEE Transactions on Intelligent Transportation Systems* 25, 2 (February 2023), 1978–1991. https://doi.org/10.1109/TITS.2023.3317637

[20] G. Harper and et al. 2019. Recycling lithium-ion batteries from electric vehicles. *Nature* 575, 7782 (2019), 75.

[21] Juan C. Herrera et al. 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies* 18, 4 (2010), 568–583.

[22] Anderson Hoke et al. 2011. Electric vehicle charge optimization including effects of lithium-ion battery degradation. In *2011 IEEE Vehicle Power and Propulsion Conference*. IEEE, 1–8.

[23] Anderson Hoke et al. 2013. Maximizing lithium ion vehicle battery life through optimized partial charging. In *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 1–5.

[24] Anderson Hoke et al. 2014. Accounting for lithium-ion battery degradation in electric vehicle charging optimization. *IEEE Journal of Emerging and Selected Topics in Power Electronics* 2, 3 (2014), 691–700.

[25] Julian Huber, David Dann, and Christof Weinhardt. 2020. Probabilistic forecasts of time and energy flexibility in battery electric vehicle charging. *Applied Energy* 262 (2020), 114525.

[26] Ryo Imai et al. 2018. Early destination prediction with spatio-temporal user behavior patterns. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–19.

[27] Andreas Janecek et al. 2015. The cellular network as a sensor: From mobile phone data to real-time road traffic monitoring. *IEEE Transactions on Intelligent Transportation Systems* 16, 5 (2015), 2551–2572.

[28] H Jannesari, MD Emami, and C Ziegler. 2011. Effect of electrolyte transport properties and variations in the morphological parameters on the variation of side reaction rate across the anode electrode and the aging of lithium ion batteries. *Journal of Power Sources* 196 (2011), 9654–9664.

[29] Peter Keil and et al. 2016. Calendar aging of lithium-ion batteries. *Journal of The Electrochemical Society* 163, 9 (2016).

[30] Ahmed S. Khwaja, Bala Venkatesh, and Alagan Anpalagan. 2021. Performance Analysis of LSTMs for Daily Individual EV Charging Behavior Prediction. *IEEE Access* 9 (2021), 154804–154814. https://doi.org/10.1109/ACCESS.2021.3128491

[31] Mohammed Khwaja, Sumer S Vaid, Sara Zannone, Gabriella M Harari, A Aldo Faisal, and Aleksandar Matic. 2019. Modeling personality vs. modeling personalidad: In-the-wild mobile data analysis in five countries suggests cultural impact on personality models. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–24.

[32] Xiangjie Kong et al. 2017. Time-location-relationship combined service recommendation based on taxi trajectory data. *IEEE Transactions on Industrial Informatics* 13, 3 (2017), 1202–1212.

[33] Gillian Lacey, Ghanim Putrus, and Edward Bentley. 2017. Smart EV charging schedules: supporting the grid and protecting battery life. *IET Electrical Systems in Transportation* 7, 1 (2017), 84–91.

[34] Tobias Lindroth, Axel Svensson, Niklas Åkerblom, Mitra Pourabdollah, and Morteza Haghir Chehreghani. 2022. An Online Learning Approach for Vehicle Usage Prediction During COVID-19. *arXiv preprint arXiv:2210.16002* (2022).

[35] Benedikt Lunz et al. 2012. Influence of plug-in hybrid electric vehicle charging strategies on charging and battery degradation costs. *Energy Policy* 46 (2012), 511–519.

[36] Bramy Pilipili Matadi et al. 2017. Effects of biphenyl polymerization on lithium deposition in commercial graphite/NMC lithium-ion pouch-cells during calendar aging at high temperature. *Journal of The Electrochemical Society* 164, 6 (2017), A1089.

[37] Joseph Molloy et al. 2023. The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland. *Transportation* 50, 5 (2023), 1983–2007.

[38] Tomás Montes et al. 2024. Degradation-conscious charge management: Comparison of different techniques to include battery degradation in Electric Vehicle Charging Optimization. *Journal of Energy Storage* 88 (2024), 111560.

[39] Gautham Ram Chandra Mouli et al. 2017. Integrated PV charging of EV fleet based on energy prices, V2G, and offer of reserves. *IEEE Transactions on Smart Grid* 10, 2 (2017), 1313–1325.

[40] Nematjon Narziev, Hwarang Goh, Kobiljon Toshnazarov, Seung Ah Lee, Kyong-Mee Chung, and Youngtae Noh. 2020. STDD: Short-term depression detection with passive sensing. *Sensors* 20, 5 (2020), 1396.

[41] Kennedy Opoku Asare, Yannik Terhorst, Julio Vega, Ella Peltonen, Eemil Lagerspetz, and Denzil Ferreira. 2021. Predicting depression from smartphone behavioral markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. *JMIR mHealth and uHealth* 9, 7 (2021), e26540.

[42] Miguel A. Ortega-Vazquez. 2014. Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty. *IET Generation, Transmission & Distribution* 8, 6 (2014), 1007–1016.

[43] Personal Data Protection Commission. 2020. Advisory Guidelines on Key Concepts in the Personal Data Protection Act. (2020). https://www.pdpc.gov.sg/guidelines-and-consultation/2020/03/advisory-guidelines-on-key-concepts-in-the-personal-data-protection-act

[44] Y. Preger and et al. 2020. Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions. *Journal of The Electrochemical Society* 167, 12 (2020), 120532.

[45] Amin Sadri et al. 2018. What will you do for the rest of the day? An approach to continuous trajectory prediction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 1–26.
[46] Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, David C Mohr, et al. 2015. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *Journal of medical Internet research* 17, 7 (2015), e4273.
[47] Muhammad Salman, Nguyen Dao, Uichin Lee, and Youngtae Noh. 2022. CSI: DeSpy: Enabling Effortless Spy Camera Detection via Passive Sensing of User Activities and Bitrate Variations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–27.
[48] Samsung Electronics. 2023. What is Adaptive Battery as new feature of Pie OS? https://www.samsung.com/za/support/mobile-devices/what-is-adaptive-battery-as-new-feature-of-pie-os/. Accessed: 2023-07-06.
[49] Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. 2011. Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*. IEEE, 166–171.
[50] M. Sedghi et al. 2015. Storage scheduling for optimal energy management in active distribution network considering load, wind, and plug-in electric vehicles uncertainties. *Journal of Renewable and Sustainable Energy* 7, 3 (2015).
[51] Sakib Shahriar et al. 2020. Machine learning approaches for EV charging behavior: A review. *IEEE Access* 8 (2020), 168980–168993.
[52] Sakib Shahriar et al. 2021. Prediction of EV charging behavior using machine learning. *IEEE Access* 9 (2021), 111576–111586.
[53] J. Song and et al. 2021. Pathways towards managing cost and degradation risk of fast charging cells with electrical and thermal controls. *Energy & Environmental Science* 14, 12 (2021), 6564–6573.
[54] NR Tague. 2004. The Quality Toolbox, ASQ (American Society for Quality).
[55] John Torous, Mathew V Kiang, Jeanette Lorme, Jukka-Pekka Onnela, et al. 2016. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. *JMIR mental health* 3, 2 (2016), e5165.
[56] Kobiljon Toshnazarov et al. 2019. EasyTrack-orchestrating large-scale mobile user experimental studies (poster). In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*.
[57] United Nations Development Programme. 2024. Personal Data Protection and Privacy Policy. https://popp.undp.org/SitePages/POPPBSModule.aspx?BPID=301&Menu=BusinessUnit
[58] Viswanath Venkatesh and Fred D Davis. 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science* 46, 2 (2000), 186–204.
[59] Huandong Wang et al. 2020. Predictability and prediction of human mobility based on application-collected location data. *IEEE Transactions on Mobile Computing* 20, 7 (2020), 2457–2472.
[60] Huandong Wang et al. 2021. Spatio-temporal urban knowledge graph enabled mobility prediction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (2021), 1–24.
[61] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 3–14.
[62] Zhe Wei, Yue Li, and Lin Cai. 2018. Electric vehicle charging scheme for a park-and-charge system considering battery degradation costs. *IEEE Transactions on Intelligent Vehicles* 3, 3 (2018), 361–373.
[63] Xuhai Xu, Prerna Chikersal, Afsaneh Doryab, Daniella K Villalba, Janine M Dutcher, Michael J Tumminia, Tim Althoff, Sheldon Cohen, Kasey G Creswell, J David Creswell, et al. 2019. Leveraging routine behavior and contextually-filtered features for depression detection among college students. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–33.
[64] Zhuo Xu. 2017. *Forecasting Electric Vehicle Arrival & Departure Time On UCSD Campus using Support Vector Machines*. Ph. D. Dissertation. UC San Diego.
[65] Elaheh Yaghoubi et al. 2024. A systematic review and meta-analysis of machine learning, deep learning, and ensemble learning approaches in predicting EV charging behavior. *Engineering Applications of Artificial Intelligence* 135 (2024), 108789.
[66] Yu Yang et al. 2020. Extending coverage of stationary sensing systems with mobile sensing systems for human mobility modeling. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–21.
[67] Yang Zhou et al. 2022. The smartphone-based person travel survey system: data collection, trip extraction, and travel mode detection. *IEEE Transactions on Intelligent Transportation Systems* 23, 12 (2022), 23399–23407.