Scheduling EV Battery Swap/Charge Operations

Jaeheon Kwak*§, Seongtae Lee†§, Kang G. Shin¶, Jinkyu Lee†‡

*School of Computing, KAIST, Republic of Korea
†Department of Computer Science and Engineering, Sungkyunkwan University, Republic of Korea
¶Department of Electrical Engineering and Computer Science, The University of Michigan, USA

***Abstract*—With the increasing popularity of electric vehicles (EVs), drivers want their vehicle batteries to be charged in a few minutes; at present, this is feasible only if battery service stations replace the EV battery pack with a fully charged battery pack. In this paper, we formulate the scheduling problem for battery swap stations, aiming to provide the drivers timing guarantees for different types of EVs, each with its own sporadic/periodic arrival pattern and deadline constraint. To solve this problem, we analyze its unique characteristics from a real-time scheduling perspective, with the main challenge being the circular timing dependency between two distinct scheduling processes: the swapping operation and the charging operation. We first derive a sufficient condition that decouples the dependency and then develop scheduling policies and timing guarantee techniques, designed for not only being specialized for the problem but also accommodating the sufficient condition in a time-predictable and resource-efficient manner. While the problem formulation and solution hold significance as the first attempt to establish real-time scheduling principles for battery swap stations, we also address how to accommodate real-world EV arrivals at a swapping station that do not necessarily follow a sporadic/periodic pattern. Finally, we evaluate the effectiveness of the proposed principles not only in addressing the formulated problem but also in accommodating real-world EV arrival patterns via simulation and a case study.**

I. Introduction

As the Electric Vehicle (EV) market grows rapidly, the inconvenience/impatience caused by the lengthy battery charging process has become increasingly evident. As a result, battery swap stations have emerged as a faster and more efficient alternative. Unlike traditional direct charging, which typically takes 1–10 hours [1], [2], the swap stations replace discharged battery packs with fully charged ones in just 3–10 minutes [3]–[5]. Many companies have recognized the potential of battery swap stations and have begun investing in and researching on them [6]–[8]. For example, NIO has already installed over 2,300 stations, with a plan for more [7]. Recent research on battery swapping (summarized in [9], [10]) has mainly focused on the location routing [11]–[16], scheduling [17]–[21], and management of wait times [22]–[26] in battery swap stations. However, little has been done to address the technical issues, despite their importance, of battery swap stations from a real-time scheduling perspective.

To fill this void, we provide a wait-time guarantee for each EV between its arrival at the battery swap station and the

§Jaeheon Kwak and Seongtae Lee are co-first authors.
‡Jinkyu Lee is the corresponding author.

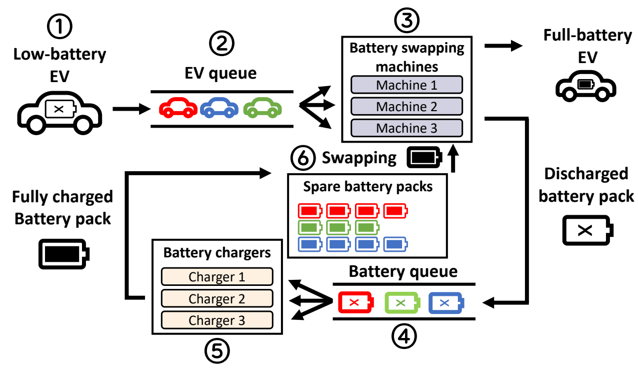


Fig. 1. Overview of a battery swap station.

completion of its battery swap. As a first step, we answer the following question:

Q1. What are the unique characteristics of the processes in the battery swap station, and how can we formulate the battery swap problem in the perspective of real-time scheduling?

The battery swap station functions as depicted in Fig. 1, and consists of the three main components: (i) battery swapping machines that replace each EV's discharged battery pack with a fully-charged one, (ii) battery chargers that fully charge a discharged battery pack taken from an EV, and (iii) spare battery packs of each type for swapping and charging batteries in parallel. When an EV arrives at the battery swap station, it waits for an available battery swapping machine. However, the EV can go through the battery swapping process only if both a battery swapping machine and a fully-charged same-type battery pack are available. On the other hand, each spare battery pack follows its own timeline, starting from the moment when an EV finishes its battery swapping. At that moment, the discharged battery pack taken from the EV is moved to the battery queue (for using battery chargers), eventually occupying a battery charger. After being fully charged, the battery pack becomes available for the next EV waiting for its discharged battery pack to be swapped out.

We observe a circular timing dependency between the swapping and charging processes, despite their contention for use of different resources (i.e., the battery swapping machines and the battery chargers, respectively). Specifically, delaying the swapping process for an EV will postpone the arrival of a returned/discharged battery pack at the battery queue, while delaying the charging process for a battery pack will defer the swapping process of a subsequent EV that has already secured an available battery swapping machine but waits for a fully-charged battery pack.

In addition to the timing dependency, one can easily observe that the completion time for both operations depends on the EV arrival pattern at the battery swap station. Analyzing real-world EV arrival patterns, we formulate the real-time scheduling problem at battery swap stations, aiming to provide wait-time guarantees to different types of EVs, each of which exhibits its own sporadic arrival pattern (with its own specified minimum inter-arrival time) and deadline constraints. Solving this problem needs to answer the following questions.

- Q2. Considering the unique characteristics, which scheduling policies are suitable for the problem?
- Q3. How do the answers to Q1 and Q2 affect the wait-time guarantees of the problem? In addition, how can we develop wait-time guarantee techniques tailored to the target problem by considering the answers?

As to Q2, we need scheduling policies for the swapping and charging processes. For the swapping process, we use the FIFO scheduling policy, which is inherently non-preemptive. Note that FIFO, albeit simple, is suitable for the swapping process, as it does not incur the additional overhead (such as moving cars) of physically altering the order of EVs for the swap service according to their non-FIFO priority. For the charging process, we design a new scheduling policy, called *quasi-non-preemptive dual-priority-FIFO scheduling*, which will be detailed in the answer of Q3.

For Q3, we exploit the answers of Q1 and Q2 as follows. First, we develop a response time analysis technique for the swapping process under an unrealistic assumption that makes it possible to remove the potential delay stemming from the charging process. Second, we derive a sufficient condition that not only relaxes the unrealistic assumption applied to the first step, but also resolves the circular timing dependency between the swapping and charging processes. This condition provides a virtual relative deadline for each battery pack arriving at the battery queue. Third, we design the quasi-non-preemptive dual-priority-FIFO scheduling policy for scheduling charging processes, which not only ensures predictability of timing of battery packs' arrival at the battery queue, but also allows for the use of the same response time analysis structure as the one developed in the first step. Then, what remains is checking if the response times for the swapping process and for the charging process are upper-bounded by the deadline constraint for each EV type and the virtual relative deadline assigned by the sufficient condition, respectively.

Since the target problem is formulated under the assumption of the minimum inter-arrival time of each type of EVs, we need strategies to relax the assumption.

- Q4. How can we accommodate real-world EV arrival patterns without assuming the minimum inter-arrival time, and which types of wait-time guarantees are achievable for the real-world patterns?

To answer Q4, we develop two BSSM (Battery Swap Station Management) policies, which enable the battery swap station to provide a deadline by which its swapping operation must be completed, even for EVs with realistic arrival patterns (not following sporadic/periodic arrivals). One is based on admission control, while the other is based on virtual arrival assignment, both of which yield different degrees and targets for timing guarantees.

To evaluate the performance of the proposed scheduling principles with their applications, we investigate (a) the timing-guarantee performance of the proposed response time analysis techniques, and (b) the run-time performance of the proposed BSSMs and scheduling policies in achieving time-predictability and resource-efficiency even without assuming the minimum inter-arrival time. The evaluation results via scheduling set generation, response time analysis implementation and run-time simulator design verify the effectiveness in terms of (a) and (b). We also conduct a case study with real-world EV arrival patterns, guiding how to utilize the proposed BSSMs while providing time-predictability.

This paper makes the following contributions.

- Formulating and solving a "first" real-time scheduling problem for battery swap stations;
- Identifying the unique characteristics of the problem from a real-time scheduling perspective (Section III);
- Applying/designing the scheduling policies for the problem (Sections IV-A and IV-C);
- Developing novel timing guarantee techniques tailored to the problem (Section IV);
- Accommodating real-world arrival patterns without assuming the minimum inter-arrival time (Section V);
- Demonstrating the effectiveness of the proposed solution via simulation and a case study (Sections VI and VII); and
- Discussing the application and extensibility of the proposed principles (Section IX).

II. System Model

We consider a battery swap station that supports $N^{\mathtt{TY}}$ different types of EVs. Each EV is equipped with its own type of battery pack, and hence we use the term "type-x" for both EV type and its battery pack type, where $1 \leq x \leq N^{\mathtt{TY}}$. We will henceforth call the battery swap station just the *station*, which consists of the following physical components.

Battery swapping machines. A battery swapping machine/chamber, whenever an EV enters the chamber, removes the discharged battery pack from the EV and replaces it with a fully-charged same-type battery pack. This swapping process substitutes the traditional direct charging process, as it is relatively fast, taking 3–10 minutes [3], [27], [28]. We consider the station is equipped with $N^{\mathtt{SW}}$ battery swapping machines. Let $C_x^{\mathtt{SW}}$ denote the actual battery swapping time of a type-x battery pack deployed in a type-x EV, which is assumed to be fixed and known.

Battery chargers. A battery charger fully charges a discharged battery pack, which was collected/returned from an EV by one of battery swapping machines. We consider the station is equipped with N^{CG} battery chargers. Let C_x^{CG} denote the worst-case time to charge a type-x battery pack from 0% to 100% SoC (State of Charge). Depending on SoC and SoH (State of Health) of a type-x battery pack to be charged, the actual charging time of a type-x battery pack varies (but does not exceed C_x^{CG}).

Spare battery packs. The station is equipped and operates with (multiple) spare battery packs for each type, in order not only to parallelize the battery swapping and charging processes, but also to handle discrepancy between the battery swapping time (taking 3–10 minutes) and the battery charging time (taking 1–10 hours). We consider the station has n_x type-x spare battery packs. Since each battery swapping process conducted in one of the battery machines replaces a discharged battery pack with a fully-charged same-type battery pack, it does not change the number of type-x spare battery packs that exist in the station, which always remains n_x. We assume $n_x \geq 1$; otherwise, each swapping machine is able to provide a fully-charged type-x battery pack, only after charging the returned pack itself, which makes the station useless. We also assume that all spare battery packs are fully charged at which the station starts its service (i.e., at the system start time).

III. Problem Statement and Analysis: Real-Time Scheduling Perspective

Wait-time guarantees for the station require the definition of two main operations associated with their queues.

Definition 1: The *swapping operation* for a type-x EV means the replacement of its type-x battery pack with a fully-charged same-type battery pack (i.e., a same-type battery pack with 100% SoC). The *EV queue* (shown in Fig. 1) means the queue for EVs that compete for use of battery swapping machines to perform the swapping operation.

Definition 2: The *charging operation* for a type-x battery pack returned by the swapping operation means charging it fully. The *battery queue* (shown in Fig. 1) means the queue for battery packs that compete for use of battery chargers to perform the charging operation.

Using Definitions 1 and 2, our timing guarantee problem is stated as follows.

G1*. We want to guarantee that the station completes the *swapping operation* for each EV within a certain time.

Note that G1* does not have any explicit timing requirement for the *charging operation*. However, since the swapping operation for a type-x EV can be performed only if a fully-charged type-x battery pack is available, the timing requirement for the swapping operation in G1* implicitly assumes that for the charging operation.

Achieving G1* is different from most existing real-time scheduling problems due to the unique characteristics of the target problem, and is involved with two scheduling processes for the swapping and charging operations. From a real-time scheduling perspective, we describe below the timelines for the two scheduling processes.

First, we present the timeline for an EV associated with the swapping operation, marked as ①–③ in Fig. 1. Once a type-x EV arrives at the station (①), it waits for the availability of a battery swapping machine in the EV queue (②). However, in order to start the swapping operation, the EV not only contends with other EVs in the EV queue, but also requires a fully-charged same-type battery pack, the latter of which depends on scheduling the charging operations to be recorded in Property 1. Once the EV occupies the resource, it finishes its swapping operation after C_x^{SW} time units (③). Then, the EV leaves the station.

Property 1: To know the completion time of the swapping operation of a type-x EV, we should know the completion time of the charging operation of the type-x battery pack to be swapped to that type-x EV, unless the number of type-x spare battery packs and the number of battery chargers are sufficiently large.

Second, the timeline for a battery pack associated with the charging operation is more complicated, marked as ①–⑥ in Fig. 1. That is, a battery pack returned by an EV goes through the same timeline as that for an EV (i.e., ①–③). Then, once its swapping operation finishes, a discharged battery pack returned by the EV arrives at the battery queue (④). The battery pack eventually secures the resource of a battery charger (⑤). Once the battery charger finishes fully charging the battery pack, it will perform the swapping operation of a subsequent EV waiting for a fully-charged battery pack (⑥). The following property is related to the timeline for a battery pack (i.e., ①–⑥) subsuming that for an EV (i.e., ①–③).

Property 2: To know the completion time of the charging operation of a type-x battery pack, we should know the completion time of the swapping operation of a type-x EV that returns that type-x battery pack.

As stated in Properties 1 and 2, there exists circular timing dependency between scheduling swapping operations and scheduling charging operations, which not only forms a unique feature of the target problem, but also makes it difficult to solve the problem.

In addition, we easily observe the following property.

Property 3: To know the completion time of the swapping and charging operations, we should have information on EV arrival patterns at the station.

According to Property 3, achieving G1* necessitates understanding EV arrival patterns at the station. To meet this requirement, we need to model the arrival patterns using real-world data. Fig. 2 shows a probability histogram of *inter-arrival times* (time intervals between EV arrivals) for two different charger models at a charging station in Palo Alto, California, USA [29]. The histogram highlights the sporadic nature of EV arrivals, with most EVs arriving after a certain inter-arrival time. Only about 1% of inter-arrival times are

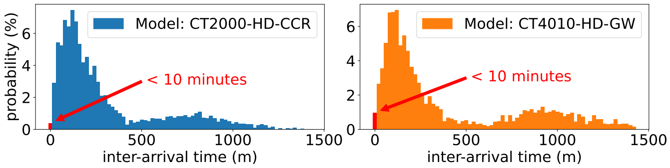


Fig. 2. Probability histogram of inter-arrival times of EVs for two different charger models at a charging station in Palo Alto.

shorter than 10 minutes for both charger models, as highlighted in Fig. 2.

Based on these observations, we consider the following sporadic arrival model for each EV type subject to timing constraints. For every $1 \leq x \leq N^{\texttt{TY}}$, type-$x$ EVs arrive at the station sporadically with at least T_x *inter-arrival time*. Once a type-x EV arrives at t, its battery pack must be replaced with a fully-charged same-type battery pack by $t + D_x^{\texttt{SW}}$, where $D_x^{\texttt{SW}}$ is the *relative deadline* of the swapping process, which is a user requirement; for example, each EV user wants the EV charging done within 30 minutes since its arrival at the station according to a survey [30]. Note that $D_x^{\texttt{SW}}$ can be smaller than, equal to, or larger than T_x. Accommodating T_x and $D_x^{\texttt{SW}}$, we can formally state the problem as:

G1. Suppose the following input parameters are given: numbers $N^{\texttt{TY}}$, $N^{\texttt{SW}}$, $N^{\texttt{CG}}$, and n_x for every $1 \leq x \leq N^{\texttt{TY}}$, and time durations T_x, $D_x^{\texttt{SW}}$, $C_x^{\texttt{SW}}$ and $C_x^{\texttt{CG}}$ for every $1 \leq x \leq N^{\texttt{TY}}$. For all $1 \leq x \leq N^{\texttt{TY}}$, we guarantee that the station completes the *swapping operation* of every type-x EV within $D_x^{\texttt{SW}}$ time units since its arrival at the station, while type-x EVs arrive at the station sporadically with at least T_x inter-arrival time.

In addition to the difficulty in achieving G1 due to the circular timing dependency between scheduling the swapping operation and scheduling charging operations, each of the two scheduling processes has a distinct relationship between the processing time ($C_x^{\texttt{SW}}$ or $C_x^{\texttt{CG}}$) and the inter-arrival time (T_x). That is, the former can be larger than the latter, which is not allowed in most traditional real-time scheduling problems. This relationship should be considered when we develop real-time scheduling techniques for our target problem.

One might argue that enforcing a minimum inter-arrival time is unrealistic. Based on the principles of Section IV that will address G1 as is, we will further extend the principles by introducing two BSSM policies in Section V to relax this assumption. We will also discuss the types of timing guarantees that the proposed BSSM policies can achieve without relying on this assumption.

IV. Development of Real-Time Scheduling Principles for Battery Swap Station

In this section, we develop real-time scheduling principles that achieve G1, based on the problem analysis in Section III. To address the issue of the circular dependency between scheduling swapping operations and scheduling charging operations, we propose the following steps.

- Step 1. We apply the most intuitive scheduling policy to the swapping operation, and assume an unrealistic configuration where the number of battery chargers and spare battery pack for each type are infinite. Under the configuration, a fully-charged battery pack is always ready to be replaced whenever a battery swapping machine is available for an EV, and therefore we focus on developing the response time analysis for the swapping operation regardless of scheduling charging operations (in Section IV-A).
- Step 2. We derive a sufficient condition that not only relaxes the unrealistic assumption applied in Step 1, but also resolves the circular timing dependency. The sufficient condition establishes a timing requirement for the charging operation, as a form of a virtual relative deadline (in Section IV-B).
- Step 3. We design a scheduling policy for the charging operation such that the virtual relative deadline derived from the sufficient condition can be utilized for scheduling charging operations in a time-predictable and resource-efficient manner. Then, we develop the response time analysis for the charging operation, in the same manner as developing that for the swapping operation (in Section IV-C).
- Step 4. We achieve G1 by testing both response time analysis for the swapping operation and the charging operation independently (in Section IV-D).

A. Scheduling and Timing Analysis of Swapping Operations under Unrealistic Configuration

For scheduling of swapping operations, we apply the following scheduling policy. Regarding prioritization, we apply the FIFO (First-In, First-Out) scheduling policy, which is the most practical for the station. Other than FIFO, the station should resolve additional overhead that physically changes the order of arriving EVs based on their priorities, which increases the cost and complexity of the station. Regarding preemptiveness, FIFO is inherently non-preemptive [31]. Note that even if other prioritization policies are applied, the non-preemptive scheduling policy is preferred for scheduling swapping operations; this is because, the battery swapping itself takes only 3–10 minutes [3], [27], [28], while the preemption requires physical movement of electric EVs from/to the battery swapping machine.

To focus on timing analysis of swapping operations regardless of scheduling charging operations, we consider the following unrealistic configuration for the station.

Situation 1: The number of battery chargers (i.e., $N^{\texttt{CG}}$) and the number of spare battery packs for each type (i.e., $\{n_x\}_{x=1}^{N^{\texttt{TY}}}$) are infinitely many.

The configuration makes it possible to perform timing analysis of the swapping operation without considering scheduling of charging operations, as a fully-charged battery pack is *always* ready to be replaced whenever a battery swapping machine is ready to serve for an EV.

Let $R_x^{\texttt{SW}}$ denote the *swapping response time* of the type-x EV, defined as an upper-bound of the time duration between

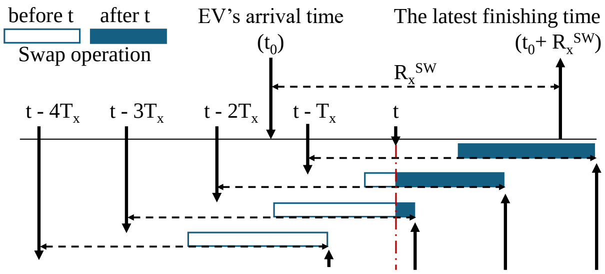


Fig. 3. Relationship between the arrival of type-x EVs and the latest time instant at which their swapping operations complete.

the arrival at the station and the completion of the swapping operation of any type-x EVs. From now on, under Situation 1, we target a type-x EV, which arrives at the station at t, and calculate its *swapping response time* $R_x^{\texttt{SW}}$, assuming $R_y^{\texttt{SW}}$ for every $1 \leq y(\neq x) \leq N^{\texttt{TY}}$ is given; we will explain how to update $R_y^{\texttt{SW}}$ in Lemma 1.

First, we calculate an upper-bound of the number of same-type (i.e., type-x) EVs that arrive at the station before t but do not finish their swapping operation until t. As shown in Fig. 3, a type-x EV arriving at the station at t_0 can finish its swapping operation no later than $t_0+R_x^{\texttt{SW}}$. Therefore, for given $R_x^{\texttt{SW}}$, the number of type-x EVs that arrive at the station before t but do not finish their swapping operation at t (by satisfying $t_0 + R_x^{\texttt{SW}} > t$) is maximized when they arrive periodically, and the number is upper-bounded by $\lfloor R_x^{\texttt{SW}}/T_x \rfloor$; for example, $\lfloor R_x^{\texttt{SW}}/T_x \rfloor = 3$ in Fig. 3. Then, the amount of back-logged (unfinished) swapping operations at t by type-x EVs which arrive before t is upper-bounded by $BL_x(R_x^{\texttt{SW}})$, which simply multiply the upper-bounded number of such type-x EVs and the battery swapping time of each EV, where

$$BL_x(L) = \begin{cases} \sum_{\alpha=1}^{\lfloor L/T_x \rfloor} C_x^{\texttt{SW}}, & \text{if } L > T_x, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Note that α (≥ 1) and L (> 0) denote the index variable and an arbitrary time interval length, respectively. If $R_x^{\texttt{SW}} \leq T_x$ holds, then the number of such type-x EVs is 0, resulting in $BL_x(R_x^{\texttt{SW}}) = 0$.

Although $BL_x(R_x^{\texttt{SW}})$ is a safe upper-bound, we can derive a tighter one. That is, some of the type-x EVs that contribute $C_x^{\texttt{SW}}$ to Eq. (1) cannot have their remaining swapping time at t as much as $C_x^{\texttt{SW}}$; for example, see two swap operations each of which has both dark blue and white portions in Fig. 3. This is because, for given $R_x^{\texttt{SW}}$, the amount of remaining swapping time at t of a type-x EV that arrives at the station at t_0 ($< t$) is upper-bounded by $t_0 + R_x^{\texttt{SW}} - t$. Therefore, instead of allowing each type-x EV to contribute $C_x^{\texttt{SW}}$ to the amount of back-logged swapping operation, we can reduce each EV's contribution, yielding $BL_x^*(R_x^{\texttt{SW}})$ as a tighter upper-bound of the amount of back-logged swapping operations at t by type-x EVs which arrive at the station before t, where

$$BL_x^*(L) = \begin{cases} \sum_{\alpha=1}^{\lfloor L/T_x \rfloor} \min\left(C_x^{\texttt{SW}}, L - \alpha \cdot T_x\right), & \text{if } L > T_x, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

As discussed in Section III, the processing time in the target problem (i.e., $C_x^{\texttt{SW}}$) can be larger than T_x, which is different from most traditional real-time scheduling problems. Due to this distinct feature, there could be multiple type-x EVs that contribute less than $C_x^{\texttt{SW}}$ to Eq. (2). Therefore, applying $BL_x^*(R_x^{\texttt{SW}})$ instead of $BL_x(R_x^{\texttt{SW}})$ results in significant improvement of the swapping response time.

Second, we calculate an upper-bound of the number of type-y ($y{\neq}x$) EVs that satisfy both cases when (i) they arrive at the station before the target type-x EV (that arrives at the station at t), thereby having a higher priority than type-x; and (ii) they do not finish their swapping operation until t. Since a later arrival at t_0 implies a later timing of $t_0+R_y^{\texttt{SW}}$, the number is maximized when type-y EVs arrive periodically. Also, since a type-y EV has a higher priority than the target type-x EV only if its arrival is no later than t according to FIFO, the upper-bounded number is calculated by $1+\lfloor R_y^{\texttt{SW}}/T_y \rfloor$, where the number 1 means the type-y EV arriving at t and the number $\lfloor R_y^{\texttt{SW}}/T_y \rfloor$ corresponds to $\lfloor R_x^{\texttt{SW}}/T_x \rfloor$ for type-x EVs arriving before t. Therefore, the amount of back-logged swapping operations at t by type-y EVs whose priority is higher than the target type-x EV is upper-bounded by $C_y^{\texttt{SW}} + BL_y^*(R_y^{\texttt{SW}})$.

Using the amount of back-logged swapping operations, we develop the swapping response time analysis.

Lemma 1: Consider a station, which satisfies Situation 1 and employs FIFO for scheduling swapping operations. Then, the swapping response time of the type-x EV is calculated by $R_x^{\texttt{SW}} = L$, where L satisfies Eq. (3).

$$L = C_x^{\texttt{SW}} + \frac{BL_x^*(L) + \sum_{y=1, y\neq x}^{N^{\texttt{TY}}} C_y^{\texttt{SW}} + BL_y^*(R_y^{\texttt{SW}})}{N^{\texttt{SW}}} \tag{3}$$

To find L that satisfies Eq. (3), for each $1 \leq x \leq N^{\texttt{TY}}$, we apply the nested iterations structure used in traditional multiprocessor backward response time analysis techniques (e.g., [32]) outlined as follows. Initially, for all $1 \leq x \leq N^{\texttt{TY}}$, $R_x^{\texttt{SW}}$ has the initial value of C_x. The structure has three levels of nested iterations: the first (the most inner), the second (the next outer), and the third (the most outer) levels.

In the first-level, a new $R_x^{\texttt{SW}}$ is obtained by updating L to the RHS (right-hand-side) of Eq. (3) until the equation is satisfied. In the second-level, after performing the first-level, if the new response time $L > D_x^{\texttt{SW}}$ holds, it is deemed *unschedulable*; otherwise, it updates $R_x^{\texttt{SW}}$ to L (when $L > R_x^{\texttt{SW}}$), or halts the second-level iteration (when $L = R_x^{\texttt{SW}}$). This second-level iteration is performed for all $1 \leq x \leq N^{\texttt{SW}}$, providing the updated $R_x^{\texttt{SW}}$ for all x. In the third level, if there is no difference in $R_x^{\texttt{SW}}$ for all $1 \leq x \leq N^{\texttt{SW}}$ compared to the previous iteration, it halts (deemed *schedulable*).

Proof: Negating Lemma 1, suppose there exists a type-x EV (called the target type-x EV) that arrives at time t and finishes the swapping operation no earlier than $t + R'$, where $R' > R_x^{\texttt{SW}}$ holds and $R_x^{\texttt{SW}}$ is calculated by the lemma. Let z denote the second term of RHS in Eq. (3) (i.e., the entire fraction) when $L = R_x^{\texttt{SW}}$. Then, by the supposition, $R_x^{\texttt{SW}} = C_x^{\texttt{SW}} + z$ holds by Eq. (3), implying $R_x^{\texttt{SW}} - C_x^{\texttt{SW}} = z$ holds.

According to the explanation for $BL_x^*(R_x^{\texttt{SW}})$ and $BL_y^*(R_y^{\texttt{SW}})$, the numerator of z is the sum of the maximum service time of the EVs whose swapping operation is guaranteed to be served before the target type-x EV. Dividing this by $N^{\texttt{SW}}$ (i.e., z itself) gives the maximum duration for which the target type-x EV waits to be served. Thus, $t+z$ is the latest time for at least one swapping machine not to service any back-logged swapping operation of EVs (excluding the target type-x EV), which arrives at the station no later than t but has unfinished swapping operation at t. Hence, if the swapping operation of the target type-x EV does not start its swapping operation until $t+z$, which satisfies $t+z = t+R_x^{\texttt{SW}} - C_x^{\texttt{SW}} < t+R' - C_x^{\texttt{SW}}$, the swapping operations of other EVs whose arrivals at the station are later than t are serviced before that of the target type-x EV, which contradicts the FIFO scheduling policy. Otherwise, the swapping operation is finished no later than $t+z+C_x^{\texttt{SW}} = t+R_x^{\texttt{SW}}$, which contradicts the supposition of the proof. ■

Using Lemma 1, we can achieve G1 under Situation 1, recorded in the following theorem.

Theorem 1: Consider a station, which satisfies Situation 1 and employs FIFO for scheduling swapping operations. Suppose $R_x^{\texttt{SW}} \le D_x^{\texttt{SW}}$ holds for every $1 \le x \le N^{\texttt{TY}}$, where $R_x^{\texttt{SW}}$ is calculated by Lemma 1. Then, G1 is achieved for the station.

Proof: By Lemma 1, the theorem holds. ■

While there has been extensive research with FIFO in the field of computer systems, e.g., [33]–[35], FIFO has received less attention in the real-time systems community, compared to other basic scheduling policies such as FP (Fixed-Priority) and EDF (Earliest Deadline First). Although there are some existing real-time scheduling studies for FIFO (e.g., [36]–[41] for uniprocessor systems or control area networks, and [31] for multiprocessor systems), no existing studies can be applied to the target problem G1, in which the processing time and the relative deadline can be larger than the inter-arrival time for multiprocessor systems. Therefore, Theorem 1 not only solves the problem for the battery swap station subject to timing constraints, but also operates as a new timing guarantee technique for FIFO scheduling on multiprocessors when the processing time and the relative deadline of each scheduling entity is larger than its inter-arrival time.

B. Decoupling Circular Timing Dependency between Swapping Operations and Charging Operations

We consider the following situation, which not only becomes more feasible than Situation 1, but also resolves the circular timing dependency between scheduling swapping operations and scheduling charging operations.

Situation 2: Whenever there exists an idle swapping machine at t at which a type-x EV has the highest priority in the EV queue, there exists a fully-charged type-x battery pack at t ready to serve for the highest-priority type-x EV.

Now, we derive a sufficient condition that yields Situation 2. Consider a type-x EV arriving at the station at t, whose discharged type-x battery pack is returned by the swapping operation at a time instant later than t; suppose that a battery charger completes the charging operation of the returned type-x battery pack no later than $t+L$ ($L>0$). This implies that a type-x battery pack returned by a type-x EV arriving at the station at t can be served for a subsequent type-x EV as long as its arrival is at $t+L$ or later. In the interval of $[t, t+L)$, there could be at most $\lceil \frac{L}{T_x} \rceil$ arrivals of type-x EVs (including the one arriving at t but excluding the one arriving at $t+L$). Therefore, as long as $n_x \ge \lceil \frac{L}{T_x} \rceil$ and the supposition hold, a fully-charged type-x battery pack is always ready until a subsequent type-x EV arrives at the station, which is a sufficient condition for Situation 2.

We derive the following lemma by translating the supposition into a virtual relative deadline of the charging operation of a type-x battery pack, denoted by $D_x^{\texttt{CG}}$.

Lemma 2: Suppose that $R_x^{\texttt{SW}} \le D_x^{\texttt{SW}}$ holds for all $1 \le x \le N^{\texttt{TY}}$ where $R_x^{\texttt{SW}}$ is calculated by Lemma 1, and $D_x^{\texttt{CG}}$ is set by Eq. (4) for all $1 \le x \le N^{\texttt{TY}}$. Also, suppose that the following statement holds for all $1 \le x \le N^{\texttt{TY}}$: every type-$x$ battery pack, discharged/returned by a type-x EV arriving at the station at t, completes its charging operation no later than $t + R_x^{\texttt{SW}} + D_x^{\texttt{CG}}$. Then, Situation 2 always holds, and therefore Lemma 1 and Theorem 1 hold without assuming Situation 1.

$$D_x^{\texttt{CG}} = n_x \cdot T_x - R_x^{\texttt{SW}} \qquad (4)$$

Proof: Negating Lemma 2, suppose there exists type-x EV for which Situation 2 does not hold.

Base case. For each of the first n_x type-x EVs, a fully-charged type-x battery is ready at the system start time (explained Section II), which satisfies Situation 2.

Inductive case. Applying Eq. (4) to the supposition of the lemma, a type-x battery pack, discharged/returned by a type-x EV arriving at the station at t (denoted by the i-th earliest arriving type-x EV), completes its charging operation no later than $t + n_x \cdot T_x$. Since $t + n_x \cdot T_x$ is the earliest arrival time of the $(i+n_x)$-th earliest arriving type-x EV, the $(i+n_x)$-th earliest-arriving type-x EV can use the battery pack returned by the i-th earliest-arriving type-x EV, where $i \ge 1$, which satisfies Situation 2.

By the base and inductive cases, Situation 2 holds for every x-type EV, which contradicts the supposition.

Once Situation 2 holds, scheduling charging operations does not affect the timing of scheduling swapping operations, making Lemma 1 and Theorem 1 hold without Situation 1. ■

C. Scheduling and Timing Analysis of Charging Operations

Thanks to Lemma 2, we can analyze the timing of the charging operation of a discharged type-x battery pack returned by a type-x EV arriving at the station at t, by transforming it into the problem of ensuring that the discharged battery pack completes the charging operation no later than $t + R_x^{\texttt{SW}} + D_x^{\texttt{CG}}$. Since $R_x^{\texttt{SW}}$ is an upper-bound, it is possible at run-time for the discharged battery pack to arrive at the battery queue no later than $t + R_x^{\texttt{SW}}$, but not necessarily at the moment. This results in unpredictability for the inter-arrival time of type-x discharged battery packs at the battery queue, which causes difficulty and pessimism of deriving an upper-bound of the amount of back-logged charging operations at a target instant, despite being essential for timing analysis. One may consider a scheduling policy that allows the discharged battery pack to intentionally defer its arrival to the battery queue until $t+R_x^{\texttt{SW}}$ (to be compared as a baseline in Section VI-C), but it is not desirable from a resource-efficiency perspective as it idles battery chargers intentionally. To address this time-unpredictability issue without compromising resource-efficiency, we propose a quasi-non-preemptive dual-priority-FIFO scheduling policy, defined as follows.

Definition 3: We define the *quasi-non-preemptive dual-priority-FIFO* scheduling policy for charging operations as follows. At any time t, a type-x battery pack belongs to a lower-priority group if $t < t_0+R_x^{\texttt{SW}}$ holds, but a higher-priority group otherwise, where t_0 is the arrival time at the station of a type-x EV that returns the type-x battery pack and $R_x^{\texttt{SW}}$ is calculated by Lemma 1. Regarding prioritization, any battery pack in the higher-priority group has a higher priority than any battery pack in the lower-priority group, and the priority of battery packs within each group is determined by FIFO. Note that the arrival time of a battery pack in the higher-priority group for applying FIFO is not the actual arrival time at the battery queue, but the time instant at which the battery pack becomes a higher-priority-group battery pack; on the other hand, the arrival time of a battery pack in the lower-priority group for applying FIFO is the actual arrival time at the battery queue. The preemption is allowed only for a higher-priority-group battery pack against a lower-priority-group battery pack.

Under quasi-non-preemptive dual-priority-FIFO scheduling, we calculate the *charging response time* of the type-x battery pack, denoted by $R_x^{\texttt{CG}}$, which is an upper-bound of the time duration between $t+R_x^{\texttt{SW}}$ and the completion of the charging operation of a type-x battery pack returned by a type-x EV arriving at the station at t. We would like to emphasize that in our definition of $R_x^{\texttt{CG}}$, the time duration does not start at the arrival time of the battery pack at the battery queue, but does start at $t+R_x^{\texttt{SW}}$.

Associated with the definition of $R_x^{\texttt{CG}}$, our design of the quasi-non-preemptive dual-priority-FIFO scheduling policy has the following advantages in terms of time-predictability (by P1 with P2), resource-efficiency (by P4), and applicability of the proposed response time analysis structure developed in Section IV-A (by P3 with P1 and P2).

- P1. If we focus on type-x battery packs belonging to the higher-priority group, time instants at which the battery packs become higher-priority-group battery packs exhibit an inter-arrival time of at least T_x (by time-predictable arrivals of higher-priority-group battery packs).
- P2. Any lower-priority-group battery packs cannot delay charging operations of any higher-priority-group battery pack (by dual-priority prioritization and preemptiveness of lower-priority-group battery packs).
- P3. The scheduling policy for higher-priority-group battery packs is exactly FIFO.
- P4. Each lower-priority-group x-type battery pack can be charged before $t_0+R_x^{\texttt{SW}}$, where t_0 is the arrival time at the station of an EV that returns the type-x battery pack.

Using the properties of the proposed scheduling policy, associated with a careful definition of $R_x^{\texttt{CG}}$, we can calculate $R_x^{\texttt{CG}}$ as follows.

Lemma 3: Consider a station, which employs quasi-non-preemptive dual-priority-FIFO scheduling for charging operations. Suppose that $R_x^{\texttt{SW}} \leq D_x^{\texttt{SW}}$ holds for every $1 \leq x \leq N^{\texttt{TY}}$, where $R_x^{\texttt{SW}}$ is calculated by Lemma 1. Then, the charging response time of the type-x EV is calculated by $R_x^{\texttt{CG}} = L$, where L satisfies Eq. (5) and $D_x^{\texttt{CG}}$ for every $1 \leq x \leq N^{\texttt{TY}}$ is set according to Eq. (4).

$$L = C_x^{\texttt{CG}} + \frac{BL_x^*(L) + \sum_{y=1, y\neq x}^{N^{\texttt{TY}}} C_y^{\texttt{CG}} + BL_y^*(R_y^{\texttt{CG}})}{N^{\texttt{CG}}} \tag{5}$$

The iteration process to find L that satisfies Eq. (5) is the same as that in Lemma 1.

Proof: When a type-x EV arrives at time t, by Lemma 1, this EV is guaranteed to return the battery pack to the station no later than $t+R_x^{\texttt{SW}}$. Suppose Lemma 3 employs an additional scheduling mechanism where, even if a battery pack is returned before $t+R_x^{\texttt{SW}}$, it is allowed to arrive at the battery queue exactly at $t+R_x^{\texttt{SW}}$. Since the minimum inter-arrival time of type-x EVs' arrivals at the station (i.e., T_x) also holds for the minimum inter-arrival time of type-x battery packs' arrivals at the battery queue, the scheduling entities in Lemma 3 with the additional scheduling mechanism are identical to the scheduling entities in Lemma 1 where EVs are replaced with battery packs. Thus, by replacing $C_x^{\texttt{SW}}$, $D_x^{\texttt{SW}}$, and $N^{\texttt{SW}}$ in Lemma 1 with $C_x^{\texttt{CG}}$, $D_x^{\texttt{CG}}$, and $N^{\texttt{CG}}$, respectively, Lemma 3 holds under the additional scheduling mechanism by Lemma 1.

When the station employs quasi-non-preemptive dual-priority-FIFO scheduling, battery packs in the higher-priority group always preempt the charging operation of battery packs in the lower-priority group, which ensures that battery packs being charged in the lower-priority group do not block the battery packs in higher-priority group. Thus, even without the additional scheduling mechanism, Lemma 3 still holds. ■

D. Achieving Timing Guarantees

Combining all principles developed in this section, we can achieve G1, recorded as follows.

Theorem 2: Consider a station, which employs FIFO for scheduling swapping operations and quasi-non-preemptive dual-priority-FIFO for scheduling charging operations. Suppose that $R_x^{\texttt{SW}} \leq D_x^{\texttt{SW}}$ and $R_x^{\texttt{CG}} \leq D_x^{\texttt{CG}}$ hold for every $1 \leq x \leq N^{\texttt{TY}}$, where $R_x^{\texttt{SW}}$ and $R_x^{\texttt{CG}}$ are calculated by Lemmas 1 and 3, respectively and $D_x^{\texttt{CG}}$ is set according to Eq. (4). Then, G1 is achieved for the station.

Proof: The theorem holds by Lemmas 1, 2 and 3. ■

V. Accommodating Real-World Arrival Patterns

In this section, we propose how to accommodate real-world EV arrival patterns that do not assume the minimum inter-arrival time, and present which types of timing guarantees are achievable without the assumption.

A. BSSM-AC: Admission Control Based Approach

To accommodate the real-world EV arrival patterns, we consider the following station management.

- M1. For a type-x EV requiring the battery swap service, the station offers a deadline by which its swapping operation must be completed, or the station rejects its entry if it is impossible to offer such a timing guarantee.
- M2. The type-x EV either accepts the offer and enters the station, or rejects it and leaves the station.
- M3. Once the type-x EV accepts the offer, the station should finish its swapping operation by the guaranteed deadline.

To utilize real-time scheduling principles developed in Section IV, we first develop a simple BSSM (Battery Swap Station Management) policy, which applies AC (Admission Control) for satisfying the inter-arrival requirement of each EV type. Let $t_x^{\texttt{prev}}(t)$ denote the latest arrival time of a type-x EV that enters the station before t.

- BSSM-AC. Consider a type-x EV that arrives at the station at t. If $t_x^{\texttt{prev}}(t)+T_x \leq t$ holds, the station grants admission for the EV to enter with $t+R_x^{\texttt{SW}}$ as a guaranteed deadline, where $R_x^{\texttt{SW}}$ is calculated by Lemma 1. Otherwise, the swap station rejects the EV to enter.

Lemma 4: Consider a station, which employs BSSM-AC for station management, FIFO for scheduling swapping operations, and quasi-non-preemptive dual-priority-FIFO for scheduling charging operations.

If the supposition of Theorem 2 holds, the station guarantees that every type-x EV entering the station at t (by receiving admission by BSSM-AC) completes its swapping operation by no later than $t+R_x^{\texttt{SW}}$, where $R_x^{\texttt{SW}}$ is calculated by Lemma 1.

Proof: It is easily observed that BSSM-AC guarantees the inter-arrival time of type-x EVs at the station is no smaller than T_x. Then, the remaining proof relies on Theorem 2. ■

B. BSSM-VA: Virtual Arrival Assignment Based Approach

Although the admission control by BSSM-AC is simple and easy to implement, it becomes highly inefficient for bursty EV arrivals. For example, imagine there is no EV arrival during 9 periods of T_x, but 10 EV arrivals within the next single period of T_x; although the average EV arrival rate during the 10 periods of T_x is one per period, the station is able to grant only one EV admission during the 10 periods. To address this issue, we develop a more resource-efficient BSSM policy, capable of accepting the same-type EVs whose inter-arrival time is less than T_x, which necessitates notions of the *virtual* arrival time and a time instant corresponding to $t_x^{\texttt{prev}}(t)$ as follows. Let $t_x^{\texttt{virt}}(t)$ denote the virtual arrival time of a type-x EV, whose actual arrival time is t. Also, let $t_x^{\texttt{prev-virt}}(t)$ (corresponding to $t_x^{\texttt{prev}}(t)$) denote the latest virtual arrival time of a type-x EV that enters the station before t. Using the notions, we develop the BSSM-VA (Virtual Arrival) policy as follows.

- BSSM-VA. Consider a type-x EV that arrives at the station at t. If $t_x^{\texttt{prev-virt}}(t) + T_x \leq t$ holds, the virtual arrival time $t_x^{\texttt{virt}}(t)$ is set to the same as the actual arrival time t. Otherwise, the virtual arrival time $t_x^{\texttt{virt}}(t)$ is set to $t_x^{\texttt{prev-virt}}(t)+T_x$. In any case, the station admits the EV to enter, with $t_x^{\texttt{virt}}(t) + R_x^{\texttt{SW}}$ as a guaranteed deadline, where $R_x^{\texttt{SW}}$ is calculated by Lemma 1.

For scheduling swapping operations with virtual arrival times, we define the FIFO-VA scheduling policy as applying the FIFO scheduling policy based on each EV's virtual arrival time (but not based on its actual arrival time). Note that while FIFO is inherently non-preemptive, we apply non-preemptiveness to FIFO-VA, which is preferred for the swapping operation.

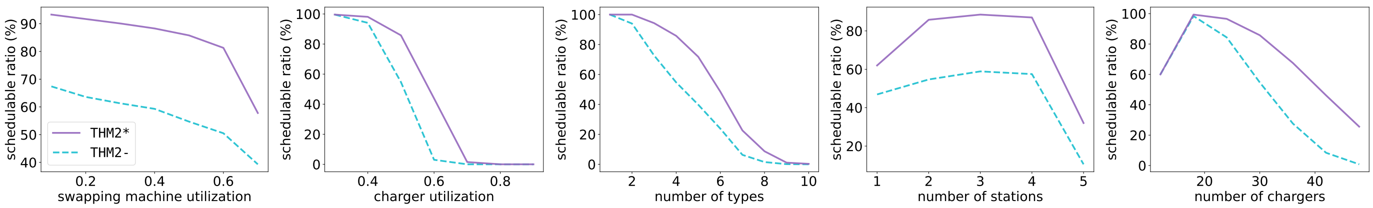
Lemma 5: Consider a station, which employs BSSM-VA for station management, FIFO-VA for scheduling swapping operations, and quasi-non-preemptive dual-priority-FIFO[1] for scheduling charging operations. If the supposition of Theorem 2 holds, the station guarantees that every type-x EV arriving at the station at t completes its swapping operation by no later than $t_x^{\texttt{virt}}(t) + R_x^{\texttt{SW}}$, where $R_x^{\texttt{SW}}$ is calculated by Lemma 1.

Proof: It is observed that BSSM-VA guarantees the virtual arrival times of any two type-x EVs are separated by at least T_x. If we regard each EV's virtual arrival time as its actual arrival time, the remaining proof relies on Theorem 2. ■

If we apply FIFO-VA instead of FIFO, there exists additional overhead that physically changes the order of arriving different-type EVs due to discrepancy between their relative order of actual arrivals and that of virtual arrivals. This is simply addressed by setting up one EV queue per each EV type; while FIFO is still valid within each queue occupied by same-type EVs, the EV with the earliest virtual arrival is selected among the most front EVs in each queue.

To summarize, in terms of the target EV coverage subject to timing guarantees, BSSM-VA includes all the EVs, whereas BSSM-AC excludes type-x EVs with inter-arrival times less than T_x; therefore, the former is more resource-efficient. In terms of the time instant by which the battery swap should be finished, BSSM-VA and BSSM-AC offer $t_x^{\texttt{virt}}(t) + R_x^{\texttt{SW}}$ and $t + R_x^{\texttt{SW}}$, respectively, when each type-x EV arrives at the station at t; therefore, the latter offers the same or tighter timing guarantees to EVs that receive admission to the station. Sections VI and VII will evaluate/discuss this tradeoff.

[1]When quasi-non-preemptive dual-priority-FIFO is applied, its criterion time instant determining the lower- or higher-priority-group should be changed from $t_0 + R_x^{\texttt{SW}}$ to $t_x^{\texttt{virt}}(t_0) + R_x^{\texttt{SW}}$, where t_0 is the arrival time at the station of an EV that returns the type-x battery pack.



(a) Varying U^{SW} (b) Varying U^{CG} (c) Varying N^{TY} (d) Varying N^{SW} (e) Varying N^{CG}

Fig. 4. Schedulable ratio with the following base parameters: $U^{\text{SW}} = 0.5$, $U^{\text{CG}} = 0.5$, $N^{\text{TY}} = 4$, $\sum_{1 \le x \le N^{\text{TY}}} n_x$=60, $N^{\text{SW}} = 2$, and $N^{\text{CG}} = 30$.

VI. Evaluation

We evaluate the proposed real-time scheduling principles and their applications. In particular, targeting battery swap stations with different parameters (to be presented in Section VI-A), we investigate the performance of (a) the proposed timing guarantee techniques in achieving G1 (in Section VI-B), and (b) the run-time behaviors of the proposed BSSMs and scheduling policies in achieving time-predictability and resource-efficiency even without assuming the minimum inter-arrival time (in Section VI-C).

A. Parameter Setting for Battery Swap Stations

We target battery swap stations each of which is equipped with multiple battery swapping machines and chargers [6], [42]. In each battery swapping machine, the swapping operation takes 3–10 minutes [3]–[5], and in each battery charger, the charging operation takes 1–10 hours (according to the NCA, NMC and LFP battery specifications [1], and the Tesla Model 3 [2]).

The parameters of each battery swap station[2] are determined within the following ranges: the number of EV types N^{TY} in $[1, 10]$, the number of swapping machines N^{SW} in $[1, 5]$, and the number of chargers N^{CG} in $[12, 48]$. Based on the actual range, the swapping time C_x^{SW} and the charging time C_x^{CG} are determined in $[3, 10]$ and $[60, 600]$ minutes, respectively. To determine some other per-type parameters, we use the utilization of swapping machines U^{SW}=$\sum_{1 \le x \le N^{\text{TY}}} C_x^{\text{SW}}/(T_x \cdot N^{\text{SW}})$ set in $[0.1, 0.7]$, and the utilization of chargers U^{CG}=$\sum_{1 \le x \le N^{\text{TY}}} C_x^{\text{CG}}/(T_x \cdot N^{\text{CG}})$ set in $[0.3, 0.9]$. The minimum inter-arrival time of each type-x EVs, T_x, is generated according to given U^{SW}, U^{CG}, and other parameters in U^{SW} and U^{CG} (except T_x). The relative deadline of the swapping operation of each EV type, D_x^{SW}, is set to 30 minutes based on the survey result in [30]. We vary the number of spare battery packs for each EV type (i.e., n_x) within $[3, 27]$.

B. Performance of Wait-Time Guarantee Techniques

To evaluate the proposed wait-time guarantee techniques, we generate scheduling sets by varying one of six parameters (i.e., U^{SW}, U^{CG}, N^{TY}, n_x, N^{SW}, and N^{CG}) according to Section VI-A, while the other five parameters are fixed as U^{SW}=0.5, U^{CG}=0.5, N^{TY}=4, $\sum_{1 \le x \le N^{\text{TY}}} n_x$=60 (thereby n_x=15 for $1 \le x \le N^{\text{TY}}$), N^{SW}=2, and N^{CG}=30 . By varying the parameters within their respective ranges, we test 45 parameter combinations. For each combination, 1,000 scheduling sets are generated, yielding 45,000 in total.

As a performance metric, we use *schedulable ratio*, defined as the ratio of the number of scheduling sets deemed schedulable (i.e., achieving G1) to the total number of tested scheduling sets. We compare two schedulability tests:

- THM2- (dashed blue line): Theorem 2 when Eqs. (3) and (5) employ $BL_x(\ell)$ in Eq. (1) instead of $BL_x^*(\ell)$ in Eq. (2), and
- THM2$*$ (solid purple line): Theorem 2 as it is, which employs $BL_x^*(\ell)$.

Fig. 4 shows the schedulable ratio according to varying U^{SW}, U^{CG}, N^{TY}, N^{SW}, and N^{CG}. THM2$*$ is shown to consistently outperform THM2 regardless of parameter variations. On average, the schedulable ratio of THM2$*$ is 38.6% higher than that of THM2-. Nevertheless, both schedulability tests exhibit similar trends in changes of their schedulable ratio according to each parameter variation, as detailed below.

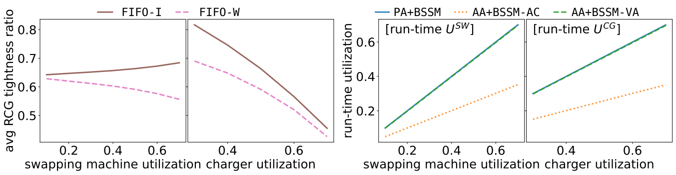
First, as depicted in Figs. 4(a) and (b), the schedulable ratio of the two tests decreases as U^{SW} and U^{CG} increase. Interestingly, the schedulable ratio decreases more steeply as U^{CG} increases (than as U^{SW} increases). This is because we set N^{CG} larger than N^{SW}, causing changes in U^{CG} to result in a greater impact on the schedulable ratio.

Second, as N^{TY} increases, the schedulable ratio drops. Fig. 4(c) depicts the schedulable ratio with varying N^{TY}, where n_x for each type is evenly distributed from $\sum_{1 \le x \le N^{\text{TY}}} n_x$=60. A large number of EV types increases the amount of backlogged operations from other types and reduces n_x, thereby decreasing the schedulable ratio.

Third, the schedulable ratio tends to be lower when N^{SW} or N^{CG} is either very small or very large, as illustrated in Figs. 4(d) and (e); we would like to remind that U^{SW} and U^{CG} are fixed in these results. Regarding N^{SW}, when it is very large, all the discharged battery packs collected by the swapping machines cannot be timely processed at the chargers, creating a bottleneck. Likewise, when N^{SW} is very small, a bottleneck occurs at the swapping machines, decreasing the schedulable ratio. The same trend is observed for N^{CG}.

In addition, we observe a positive correlation between the schedulable ratio and n_x (not shown in the figure). The more battery packs in the station, the longer D_x^{CG} according to Eq. (4), which enhances the schedulable ratio.

[2]Considering the analogy between a set of diverse station parameters and a set of real-time tasks with varying parameters, we will henceforth call a battery swap station under evaluation a *scheduling set*.



(a) Average RCG tightness ratio (b) Run-time utilization

Fig. 5. Run-time performance according to policy and arrival pattern.

C. Run-time Performance

Focusing on the run-time behaviors of the proposed principles and their applications, we now evaluate (RB1) the resource-efficiency of the proposed scheduling policy for the charging operation, and (RB2) the capability of the proposed BSSMs in granting admission to more EVs, even when the actual inter-arrival time of type-x EVs is less than T_x.

For this, we develop a battery swap station simulator, which emulates the entire process of the station, from the arrival of EVs to the scheduling of battery swapping machines and chargers. The simulator considers two arrival models for each type-x EV: (`PA` or Periodic Arrival) a fixed inter-arrival time of T_x, and (`AA` or Aperiodic Arrival) an aperiodic inter-arrival time uniformly distributed in $[0.5{\cdot}T_x, 1.5{\cdot}T_x]$ whose average is T_x; under `AA`, the actual inter-arrival time of type-x EVs could thus be less than T_x. We conduct simulations by varying the parameters $U^{\texttt{SW}}$ and $U^{\texttt{CG}}$ with other parameters set to the fixed values applied in Section VI-B. Out of the generated 1,000 scheduling sets for each pair of $U^{\texttt{SW}}$ and $U^{\texttt{CG}}$, we simulate the sets guaranteed to be schedulable by `THM2*`. Each scheduling set is simulated for 30 days (43,200 minutes) with a time granularity of 0.1 minute.

To evaluate RB1, we target two scheduling policies for the charging operation: (`FIFO-W`, *ours*) quasi-non-preemptive dual-priority-FIFO, which is work-conserving, and (`FIFO-I`, *baseline*) a FIFO scheduling policy that makes a discharged battery pack intentionally defer its arrival to the battery queue until $t + R_x^{\texttt{SW}}$, which yields idling of battery chargers, where t is the arrival time of a type-x EV that returned that discharged battery pack. We compare `FIFO-W` and `FIFO-I`, when they are applied to the arrival pattern of `AA` (shown in Fig. 5(a)).

As to RB1, we measure the average RCG tightness ratio for a type-x battery pack, defined as the ratio of the run-time $R_x^{\texttt{CG}}$ by simulation[3] to the calculated $R_x^{\texttt{CG}}$ by Lemma 3. Fig. 5(a) illustrates the average RCG tightness ratio when varying $U^{\texttt{SW}}$ and $U^{\texttt{CG}}$. We identify that the average RCG tightness ratio of `FIFO-W` is on average 10.2% lower than that of `FIFO-I`, where a lower average RCG tightness ratio implies earlier charging completion. This holds because, `FIFO-I` disallows charging during the period between the completion of the swap operation (that returns the battery pack) and the criterion time instant defined in Definition 3, while `FIFO-W` may utilize an idle charger during this period. Therefore, `FIFO-W` achieves resource-efficiency by its work-conserving nature while providing time-predictability to its wait-time guarantee technique.

To evaluate RB2, we compare the following pairs of a BSSM policy and an arrival pattern: `PA+BSSM`, `AA+BSSM-AC` and `AA+BSSM-VA`, where `X+Y` means applying the BSSM of `Y` to the arrival pattern of `X`; note that BSSM-AC is equivalent to BSSM-VA under `PA`, yielding unified `PA+BSSM`.

Regarding RB2, we measure the run-time utilization of battery swapping machines and chargers through simulations, as depicted in Fig. 5(b). The run-time utilization $U^{\texttt{SW}}$ (*likewise* $U^{\texttt{CG}}$) is the ratio of the time swapping machines (*likewise* chargers) are utilized to the total simulation time, indicating the number of EVs admitted to the station. When EVs periodically arrive (i.e., `PA+BSSM`), any BSSM always grants their admission, and the run-time utilization of battery swapping machines and chargers follow $U^{\texttt{SW}}$ and $U^{\texttt{CG}}$. On the other hand, when EVs arrive aperiodically (`AA+BSSM-`$\star$), BSSMs have to control their admissions. `AA+BSSM-AC` fails to fully utilize the battery swapping machines and chargers as it rejects all EVs arrived earlier than $t_x^{\texttt{prev}}(t){+}T_x$. In contrast, `AA+BSSM-VA` effectively manages early-arrived EVs, achieving an actual utilization level nearly identical to the `PA+BSSM` as shown in Fig. 5(b), at the expense of delaying when the battery swap is guaranteed to be finished (from $t{+}R_x^{\texttt{SW}}$, to $t_x^{\texttt{virt}}(t){+}R_x^{\texttt{SW}}$). This tradeoff is related to modeling real-world arrival patterns, to be discussed in the next section.

VII. Case Study with Real-World Arrival Patterns

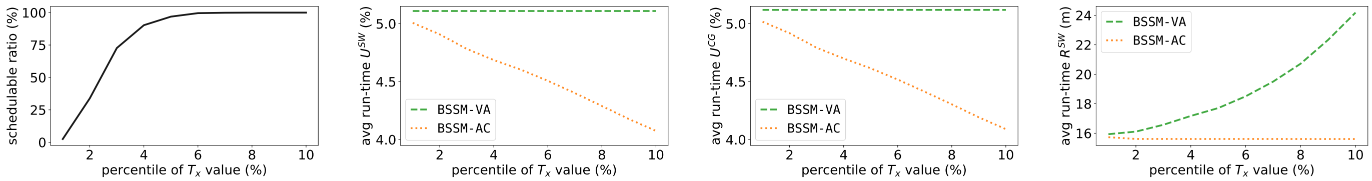
We conduct a case study applying the proposed scheduling principles to real-world arrival patterns, in order to evaluate their effectiveness in practical settings and demonstrate key considerations for arrival pattern modeling.

The case study targets the arrival pattern of EVs visited at the MPL #6 charging station in the Palo Alto area in California, USA [29]. This charging station has two charger models, with 2,601 and 4,000 arrival records for each model, shown in left and right sub-figures in Fig. 2, respectively. We split these arrival records in half, resulting in four different arrival patterns for generating scheduling sets.

In this case study, it is important to guide how to model real-world arrival patterns, which is determined by the choice of T_x in our problem statement in G1. To this end, we generate 1,000 scheduling sets using the same method as described in Sections VI-A and VI-B, and then expand these sets by varying their T_x from the bottom 1%, 2%, 3%, …, 10% of each actual inter-arrival pattern (lower percentiles mean shorter T_x), resulting in a total of 10,000 scheduling sets.

First, we examine how the choice of T_x changes schedulability when all other factors remain constant. We perform schedulability tests by varying T_x, as shown in Fig. 6(a). The schedulable ratio is low at 2.5% when T_x is set to the 1st percentile. However, as T_x increases, the ratio also increases, reaching 96.9% when T_x is set to the 5th percentile. This result

[3]The run-time $R_x^{\texttt{CG}}$ means the battery pack's time duration between its arrival at the charger and the completion of the charging operation. Likewise, the run-time $R_x^{\texttt{SW}}$ means the EV's time duration between its actual arrival at the station and the completion of the swapping operation.



(a) Schedulable ratio (b) Run-time swapping mach. util. (c) Run-time charger util. (d) Average run-time R^{SW}

Fig. 6. Schedulable ratio and run-time performance under different BSSMs and T_x assignment with real-world arrival patterns.

indicates that a larger T_x allows more scheduling sets to meet the timing constraints.

Second, to evaluate the effectiveness of our scheduling policies and the impact of T_x choice, our scheduling policies are applied to real-world arrival patterns. Figs. 6(b) and (c) illustrate the run-time U^{SW} and U^{CG}, and Fig. 6(d) depicts the average run-time R^{SW} when using BSSM-AC and BSSM-VA. While our BSSM-AC and BSSM-VA successfully manage the station by keeping timing constraints with real-world EV arrival patterns, there are differences in runtime performance depending on the management strategy employed.

On the one hand, BSSM-AC reduces the utilization of both the swapping machines and chargers, and this effect becomes more significant as T_x increases. For example, when using BSSM-AC, the average run-time U^{SW} (swapping machine utilization) decreases from 5.0% to 4.1% as T_x increases from the 1st to 10th percentile as shown in Fig. 6(b). The same trend appears in average run-time U^{CG} (charger utilization), as depicted in Fig. 6(c). This occurs as BSSM-AC rejects more EV arrivals with their inter-arrival times shorter than T_x.

On the other hand, BSSM-VA increases the completion time of swapping operations, and this trend also intensifies as T_x increases. Since BSSM-VA virtually delays EV arrivals to their virtual arrivals, it shows a larger run-time R^{SW} than BSSM-AC. For instance, the average run-time R^{SW} for BSSM-VA increases from 15.9 minutes to 24.2 minutes as T_x rises from 1% to 10%, as shown in Fig. 6(d).

These results suggest that station management requires the appropriate selection of T_x and the BSSM for a given setting. That is, too small T_x may lead to the failure of timing guarantees, while too large T_x entails a decrease of swapping machine and charger utilization (for BSSM-AC) or a delay of swapping completion (for BSSM-VA). Therefore, the station operator determines the minimum value of T_x according to the timing guarantee (for both), and sets its maximum value based on the station's resource utilization (for BSSM-AC) or EV users' timing requirements (for BSSM-VA).

VIII. Related Work

Due to its capability to provide a fully charged battery pack within a few minutes, battery swap stations have been widely researched, which can be categorized as follows.

Battery swap station location–routing problem. This research area focuses on finding the optimal placement of battery swap stations and the routing of EVs. Yang and Sun introduced the integration of location and routing problem for battery swap stations [11]. Hof *et al.* improved solutions for the problem using the AVNS algorithm [12]. Moon *et al.* proposed multiple methods for determining the location of battery swap stations targeting the e-Bus system [13]. These studies highlight potential cost savings and environmental benefits [14], [15] without addressing wait-time guarantees.

Scheduling of operations in battery swap stations. Extensive research has been done on scheduling the charging and swapping of battery packs at battery swap stations. Yang *et al.* proposed a scheduling models that prioritize battery packs for swapping, considering battery characteristics [17]. As for the scheduling of charging, many studies mainly focused on minimizing electricity costs [18]–[21]. However, none of them provided wait-time guarantees.

Management of wait times at battery swap stations. Several studies considered the wait time of EV users. Some studies aimed to maximize the quality of service (QoS) metric that includes the wait time for battery swapping [22], [23]. Other approaches directly reduced users' wait time for battery swapping by controlling the charging process [24] or enabling reservation [25], [26]. While effectively dropping the average wait time, existing studies cannot guarantee the deadlines of completing the swapping or charging battery packs.

Approaches to reduce waiting time. Numerous studies have investigated scheduling problems with the objective of minimizing waiting times across various industrial fields. For example, Chu *et al.* [43] focused on reducing the total waiting time of successive jobs, while Ye *et al.* [44] aimed to minimize waiting time variance on a single resource. Yu *et al.* [45] investigated how to reduce waiting time variance in wafer fabrication. In the EV domain, Hussain *et al.* [46] formulated the optimization of EV waiting times in public electric vehicle supply equipment as a fuzzy integer linear programming problem. While these approaches might seem applicable to battery swap stations, they did not address two critical factors: the circular dependency inherent in battery swapping operations and the consideration of timing guarantees.

In summary, while there were many interesting studies for battery swap stations, our work is the first to establish real-time scheduling principles for battery swap stations, which not only solves a real-time scheduling problem, but also accommodates realistic EV arrival patterns.

IX. Discussion

The principles and framework proposed in this paper can be applied and extended further as follows.

Application to station installation. Our principles facilitate cost optimization in the installation of battery swap stations. By applying the modeled EV arrival patterns and setting a target deadline for the swapping operation, we can formulate the problem of finding a minimum-cost configuration for the number of battery packs, swapping machines and chargers. Solving the problem does not require extensive simulations, as each configuration can be evaluated by checking Theorem 2.

Extension to other domains. Our approaches to resolving circular timing dependencies in battery swap stations can be extended to other domains with similar characteristics. For example, our BSSMs and/or scheduling policies could be effectively applied to pallet management in factory production lines, where pallets exhibit circular timing dependencies.

Leveraging run-time flexibility for advanced functionalities. We used C_x^{CG} to account for the worst-case time required to charge a type-x battery pack from 0% to 100% SoC. In practice, however, the returned battery pack may have non-zero SoC, and some EV users might accept a battery pack charged to 80% rather than fully charged. Although such optimistic behaviors cannot be applied from a worst-case perspective, they introduce additional run-time flexibility, to be incorporated into our framework to enable additional functionalities, such as (i) managing the thermal risk, aging and/or SoH of battery packs, (ii) enabling station reservations for EV users, and (iii) reducing the electricity cost of charger usage by accounting for time-varying electricity rates.

X. Conclusion

In this paper, we formulated a real-time scheduling problem for battery swap stations, analyzed the problem from a timing-guarantee perspective, and developed novel solutions, all of which are the first to establish real-time scheduling principles for battery swap stations. In addition, we proposed strategies to utilize the proposed principles for accommodating real-world EV arrival patterns, and validated by our evaluation results with the patterns. In future, we plan to address the three issues raised in Section IX.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00438248, RS-2024-00347516). It was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2024-00451982). Kang Shin was supported in part by the US Department of Energy, ARPA-E, under Grant DE-AR0001975.

References

[1] M. Koniak and A. Czerepicki, "Selection of the battery pack parameters for an electric vehicle based on performance requirements," in *IOP conference series: Materials science and engineering*, vol. 211, no. 1. IOP Publishing, 2017.

[2] "Tesla Model 3," https://ev-database.org/car/1991/Tesla-Model-3.

[3] H. Wu, "A survey of battery swapping stations for electric vehicles: Operation modes and decision scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10 163–10 185, 2021.

[4] "NIO Power," https://www.nio.com/nio-power?

[5] I. Boudway, "A california startup now offers a full ev battery in just 10 minutes," *bloomberg*, 2021, https://www.bloomberg.com/news/articles/2021-06-10/california-startup-ample-offers-a-full-ev-battery-in-just-10-minutes.

[6] S. Doll, "Ample and ENEOS are bringing the first-ever EV battery swap stations to Japan," *electrek*, 2024, https://electrek.co/2024/03/28/ample-eneos-bringing-first-ev-battery-swap-stations-to-japan/.

[7] "NIO Joins Hands With CATL in Innovation of Long-Life Batteries," 2024, https://www.nio.com/news/NIO-Joins-Hands-With-CATL.

[8] GlobeNewswire, "Battery Swapping for Electric Vehicles Market Size Will Hit US$16 Bn Mark by 2030," *Fairfield Consultancy Services OPC Pvt Ltd*, 2024, https://www.globenewswire.com/news-release/2023/08/02/2717245/0/en/Battery-Swapping-for-Electric-Vehicles-Market-Size-Will-Hit-US-16-Bn-Mark-by-2030-Fairfield-Market-Research.html.

[9] D. Cui, Z. Wang, P. Liu, S. Wang, D. G. Dorrell, X. Li, and W. Zhan, "Operation optimization approaches of electric vehicle battery swapping and charging station: A literature review," *Energy*, vol. 263, p. 126095, 2023.

[10] L. Zhong and M. Pei, "Optimal design for a shared swap charging system considering the electric vehicle battery charging rate," *Energies*, vol. 13, no. 5, p. 1213, 2020.

[11] J. Yang and H. Sun, "Battery swap station location-routing problem with capacitated electric vehicles," *Computers & operations research*, vol. 55, pp. 217–232, 2015.

[12] J. Hof, M. Schneider, and D. Goeke, "Solving the battery swap station location-routing problem with capacitated electric vehicles using an avns algorithm for vehicle-routing problems with intermediate stops," *Transportation research part B: methodological*, vol. 97, pp. 102–112, 2017.

[13] J. Moon, Y. J. Kim, T. Cheong, and S. H. Song, "Locating battery swapping stations for a smart e-bus system," *Sustainability*, vol. 12, no. 3, p. 1142, 2020.

[14] S. Kim, R. E. Pelton, T. M. Smith, J. Lee, J. Jeon, and K. Suh, "Environmental implications of the national power roadmap with policy directives for battery electric vehicles (bevs)," *Sustainability*, vol. 11, no. 23, p. 6657, 2019.

[15] Z. Chao and C. Xiaohong, "Optimizing battery electric bus transit vehicle scheduling with battery exchanging: Model and case study," *Procedia-Social and Behavioral Sciences*, vol. 96, pp. 2725–2736, 2013.

[16] S. S. Amiri, S. Jadid, and H. Saboori, "Multi-objective optimum charging management of electric vehicles through battery swapping stations," *Energy*, vol. 165, pp. 549–562, 2018.

[17] J. Yang, W. Liu, K. Ma, Z. Yue, A. Zhu, and S. Guo, "An optimal battery allocation model for battery swapping station of electric vehicles," *Energy*, vol. 272, p. 127109, 2023.

[18] W. Zhan, Z. Wang, L. Zhang, P. Liu, D. Cui, and D. G. Dorrell, "A review of siting, sizing, optimal scheduling, and cost-benefit analysis for battery swapping stations," *Energy*, vol. 258, p. 124723, 2022.

[19] Y. Gao, J. Yang, M. Yang, and Z. Li, "Deep reinforcement learning based optimal schedule for a battery swapping station considering uncertainties," *IEEE Transactions on Industry Applications*, vol. 56, no. 5, pp. 5775–5784, 2020.

[20] M. R. Sarker, H. Pandžić, and M. A. Ortega-Vazquez, "Optimal operation and services scheduling for an electric vehicle battery swapping station," *IEEE transactions on power systems*, vol. 30, no. 2, pp. 901–910, 2014.

[21] Y. Li, Z. Yang, G. Li, Y. Mu, D. Zhao, C. Chen, and B. Shen, "Optimal scheduling of isolated microgrid with an electric vehicle battery swapping station in multi-stakeholder scenarios: A bi-level programming approach via real-time pricing," *Applied energy*, vol. 232, pp. 54–68, 2018.

[22] B. Sun, X. Tan, and D. H. Tsang, "Optimal charging operation of battery swapping stations with qos guarantee," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2014, pp. 13–18.

[23] Z. Ding, W. Tan, W. Lu, and W.-J. Lee, "Quality-of-service aware battery swapping navigation and pricing for autonomous mobility-on-demand system," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 8247–8257, 2022.

[24] Z. Moghaddam, I. Ahmad, D. Habibi, and Q. V. Phung, "Smart charging strategy for electric vehicle charging stations," *IEEE Transactions on transportation electrification*, vol. 4, no. 1, pp. 76–88, 2017.

[25] X. Zhang, Y. Cao, L. Peng, N. Ahmad, and L. Xu, "Towards efficient battery swapping service operation under battery heterogeneity," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6107–6118, 2020.

[26] Y. Cao, X. Zhang, B. Zhou, X. Duan, D. Tian, and X. Dai, "Mec intelligence driven electro-mobility management for battery switch service," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4016–4029, 2020.

[27] E. Hai, "10 Billion Kilometers of Driving Reveals NIO's Users Preference in Battery Swapping," 2023, https://www.nio.com/news/10-Billion-Kilometers.

[28] C. Crownhart, "How 5-minute battery swaps could get more EVs on the road," *MIT Technology Review*, 2023, https://www.technologyreview.com/2023/05/17/1073265/how-5-minute-battery-swaps-could-get-more-evs-on-the-road/.

[29] C. of Palo Alto, "Electric vehicle charging station usage," 2021, https://data.cityofpaloalto.org/dataviews/257812/electric-vehicle-charging-station-usage-july-2011-dec-2020/.

[30] S. G. Mobility, "Affordability tops charging and range concerns in slowing EV demand," *S&P Global*, 2023, https://www.spglobal.com/mobility/en/research-analysis/affordability-tops-charging-and-range-concerns-in-slowing-ev-d.html.

[31] H. Leontyev and J. H. Anderson, "Tardiness bounds for FIFO scheduling on multiprocessors," in *Proceedings of Euromicro Conference on Real-Time Systems (ECRTS)*, 2007.

[32] S. Baruah, M. Bertogna, and G. Buttazzo, *Multiprocessor scheduling for real-time systems*. Springer, 2015.

[33] J. Li, T. Ma, M. Tang, W. Shen, and Y. Jin, "Improved fifo scheduling algorithm based on fuzzy clustering in cloud computing," *Information*, vol. 8, no. 1, p. 25, 2017.

[34] J. Qian, W. Srisa-An, S. Seth, H. Jiang, D. Li, and P. Yi, "Exploiting fifo scheduler to improve parallel garbage collection performance," in *Proceedings of the 12th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments*, 2016, pp. 109–121.

[35] O. Beaumont, L. Marchal, V. Rehn, and Y. Robert, "FIFO scheduling of divisible loads with return messages under the one-port model," in *Proceedings of the 20th IEEE International Parallel & Distributed Processing Symposium*, 2006.

[36] L. George and P. Minet, "A FIFO worst case analysis for a hard real-time distributed problem with consistency constraints," in *Proceedings of International Conference on Distributed Computing Systems*, 1997, pp. 441–448.

[37] M. Nasri, R. I. Davis, and B. B. Brandenburg, "FIFO with offests: high schedulability with low overheads," in *Proceedings of IEEE Real-Time Technology and Applications Symposium (RTAS)*, 2018.

[38] K. Bedarkar, M. Vardishvili, S. Bozhko, M. Maida, and B. B. Brandenburg, "From intuition to coq: a case study in verified response-time analysis of FIFO scheduling," in *Proceedings of IEEE Real-Time Systems Symposium (RTSS)*, 2022.

[39] R. I. Davis, S. Kollmann, V. Pollex, and F. Slomka, "Controller area network (can) schedulability analysis with fifo queues," in *Proceedings of Euromicro Conference on Real-Time Systems (ECRTS)*, 2011, pp. 45–56.

[40] R. I. Davis and N. Navet, "Controller area network (can) schedulability analysis for messages with arbitrary deadlines in fifo and work-conserving queues," in *the 9th IEEE International Workshop on Factory Communication Systems*, 2012, pp. 33–42.

[41] R. I. Davis, S. Kollmann, V. Pollex, and F. Slomka, "Schedulability analysis for controller area network (can) with fifo queues priority queues and gateways," *Real-Time Systems*, vol. 49, pp. 73–116, 2013.

[42] W. Li, Y. Li, H. Deng, and L. Bao, "Planning of electric public transport system under battery swap mode," *Sustainability*, vol. 10, no. 7, p. 2528, 2018.

[43] C. Chu and M.-C. Portmann, "Job-shop scheduling to minimize total waiting time," *Applied stochastic models and data analysis*, vol. 9, no. 2, pp. 177–185, 1993.

[44] N. Ye, X. Li, T. Farley, and X. Xu, "Job scheduling methods for reducing waiting time variance," *Computers & Operations Research*, vol. 34, no. 10, pp. 3069–3083, 2007.

[45] T.-S. Yu, H.-J. Kim, and T.-E. Lee, "Minimization of waiting time variation in a generalized two-machine flowshop with waiting time constraints and skipping jobs," *IEEE Transactions on Semiconductor Manufacturing*, vol. 30, no. 2, pp. 155–165, 2017.

[46] S. Hussain, Y.-S. Kim, S. Thakur, and J. G. Breslin, "Optimization of waiting time for electric vehicles using a fuzzy inference system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15 396–15 407, 2022.