Learning-Augmented Algorithms for the Bahncard Problem

Hailiang Zhao[†,‡], Xueyan Tang[‡], Peng Chen[†], and Shuiguang Deng[†]

[†]Zhejiang University
[‡]Nanyang Technological University

November 11, 2024

Outline

Outline

Introduction
- The Bahncard Problem
- Learning-Augmented Algorithms
- Our Results

Preliminaries

Algorithm with Predictions
- SUM_w
- FSUM
- PFSUM

Experiments

The Bahncard Problem

The Bahncard is a railway pass of the German railway company, which provides a discount on all train tickets for a fixed time period of pass validity.[1]

When a travel request arises, a traveler can buy the train ticket with the regular price, or purchase a Bahncard first and get entitled to a discount on all train tickets within its valid time.

The Bahncard problem is *an online cost minimization problem*, whose objective is to minimize the overall cost of pass and ticket purchases, without knowledge of future travel requests.

[1] https://www.bahn.com/en/offers/bahncard

Competitive Ratio

For online algorithms, we use *competitive ratio* to analyze their performance.

Competitive Ratio

Competitive ratio is the worst-case ratio across all inputs between the costs of the online algorithm and an optimal offline one:

$$\mathtt{CR}_{\mathsf{ALG}} = \max_{\forall I} \frac{\mathsf{ALG}(I)}{\mathsf{OPT}(I)}, \tag{1}$$

where I is an arrival instance of the given problem. $\mathsf{ALG}(I)$ and $\mathsf{OPT}(I)$ are respectively the cost of ALG and OPT incurred in I.

Extensions of the Bahncard Problem

The Bahncard problem reveals a *recurring renting-or-buying* phenomenon, where an online algorithm needs to *irrevocably* and *repeatedly* decide between

1. a cheap short-term solution, and
2. an expensive long-term one

with an unknown future.

We can find its applications in numerous applications in computer systems. To name a few:

- Cloud reservation (reserved instances and spot instances)
- Edge caching (caching contents at the network edge with extra cost)
- Refactoring versus working with a poor design
- ...

BP(β, C, T) and the SUM Algorithm

The Bahncard problem is instantiated by three parameters and denoted by BP(β, C, T), meaning that a Bahncard costs C, reduces any ticket price p to βp for some $\beta \in [0, 1)$, and is valid for a time period of T.

Fleischer was the first to study the Bahncard problem [1]. By extending the optimal 2-competitive break-even algorithm for ski-rental, Fleischer proposed an optimal deterministic online algorithm named SUM for BP(C, β, T), which is $(2 - \beta)$-competitive.[2]

[2]We will show how SUM works later.

Learning-Augmented Algorithms

Learning-augmented algorithms aim to leverage machine-learned predictions to improve the performance in both theory and practice [2]. Algorithms with *possibly imperfect* predictions, whose performance is measured with respect to *the quality of predictions*, have found applications in numerous important problems [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

For a comprehensive collection of the literature, see `https://algorithms-with-predictions.github.io/`.

Consistency and Robustness

Consistency and Robustness

Consistency refers to the competitive ratio under perfect predictions while robustness refers to the upper bound of the competitive ratio when the predictions can be arbitrarily bad.

Following the definition in [14, 3], we take the competitive ratio of a learning-augmented online algorithm ALG as a function $\mathtt{CR}_{\mathsf{ALG}}(\eta)$ of the prediction error η. ALG is δ-*consistent* if $\mathtt{CR}_{\mathsf{ALG}}(0) = \delta$, and ϑ-*robust* if $\mathtt{CR}_{\mathsf{ALG}}(\eta) \leq \vartheta$ for all η.

A Recent Work on the Bahncard Problem

Bamas et al. designed an algorithm, PDLA, for BP(β, C, T), that is $\lambda/(1-\beta+\lambda\beta)\cdot(e^\lambda-\beta)/(e^\lambda-1)$-consistent and $(e^\lambda-\beta)/(e^\lambda-1)$-robust when $C\to\infty$ (in which case the optimal solution is to never buy any Bahncard), where $\lambda\in(0,1]$ is a hyper-parameter [15].[3]

Their algorithm is built on the online primal-dual framework proposed in [16]. Notably, their algorithm demands a complete solution as input advice, which implies a predicted *complete* sequence of travel requests over an arbitrarily long timespan, which is impractical for real employment.

[3] Their method is limited to scenarios with *slotted* time and *uniform* ticket prices for all travel requests. BTW, their result is quite weak.

Our Results

We develop an algorithm, named PFSUM, which takes short-term predictions on future trips as inputs.

PFSUM

At any time t when a travel request arises and there is no a valid Bahncard, a prediction on the total (regular) ticket price of all travel requests in the upcoming interval $[t, t+T)$ is made. Incorporating this prediction, PFSUM purchases a Bahncard at time t when

1. the total ticket price in the past interval $(t-T, t]$ is at least γ and
2. the predicted total ticket price in $[t, t+T)$ is also at least γ, where

$$\gamma := \frac{C}{1-\beta}. \tag{2}$$

Our Results

Denoting by η the maximum prediction error, we derive that

$$\mathtt{CR}_{\mathsf{PFSUM}}(\eta) = \begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \le \eta \le \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{3}$$

The result shows that PFSUM is $2/(1+\beta)$-consistent and $1/\beta$-robust, and its competitive ratio degrades smoothly as the prediction error increases.

We are the first to present the competitive ratio of the designed algorithm (for the Bahncard problem) with any given prediction error, rather than that with only endpoints ($\eta = 0$ and $\eta \to \infty$).

Outline

Problem Restated

BP(C, β, T) is a request-answer game between an online algorithm ALG and an *adversary*. The adversary presents a finite sequence of travel requests $\sigma = \sigma_1 \sigma_2 \cdots$, where each σ_i is a tuple (t_i, p_i) that contains the travel time $t_i \geq 0$ and the regular ticket price $p_i \geq 0$. The travel requests are presented in chronological order: $0 \leq t_1 < t_2 < \cdots$.

ALG needs to react to each travel request σ_i. If ALG does not have a valid Bahncard, it can opt to buy the ticket with the regular price p_i, or first purchase a Bahncard which costs C, and then pay the ticket price with a β-discount, i.e., βp_i. A Bahncard purchased at time t is valid during the time interval $[t, t + T)$.

Algorithm Cost and Competitive Ratio

We say σ_i is a *reduced request* of ALG if ALG has a valid Bahncard at time t_i. Otherwise, σ_i is a *regular request* of ALG. We use $\mathsf{ALG}(\sigma_i)$ to denote ALG's cost on σ_i:

$$\mathsf{ALG}(\sigma_i) = \begin{cases} \beta p_i & \text{ALG has a valid Bahncard at } t_i, \\ p_i & \text{otherwise.} \end{cases} \tag{4}$$

We denote by $\mathsf{ALG}(\sigma)$ the total cost of ALG for reacting to all the travel requests in σ. The competitive ratio of ALG can be formally defined by $\mathtt{CR}_{\mathsf{ALG}} := \max_\sigma \mathsf{ALG}(\sigma)/\mathsf{OPT}(\sigma)$.

Partial Cost Incurred in an Interval

We use $\mathsf{ALG}(\sigma;\mathcal{I})$ to denote the partial cost incurred during a time interval $\mathcal{I}$: $\mathsf{ALG}(\sigma;\mathcal{I}) = C \cdot x + \sum_{i:t_i\in\mathcal{I}} \mathsf{ALG}(\sigma_i)$, where x is the number of Bahncards purchased by ALG in $\mathcal{I}$. Additionally, we use $c(\sigma;\mathcal{I})$ to denote the total regular cost in $\mathcal{I}$: $c(\sigma;\mathcal{I}) := \sum_{i:t_i\in\mathcal{I}} p_i$.

Given a time length l, we define the *l-recent-cost* of σ at time t as $c(\sigma;(t-l,t])$. Similarly, we define the *l-future-cost* of σ at time t as $c(\sigma;[t,t+l))$.

Predicted Future Cost and Regular Recent Cost

When a travel request (t, p) arises, a short-term prediction of the total regular cost in the time interval $[t, t + T)$ can be made.

To represent prediction errors, we use $\hat{c}(\sigma; [t, t + T))$ to denote the predicted total regular cost in $[t, t + T)$. Sometimes we are concerned about the regular requests of an algorithm ALG in a recent time interval. Thus, we further define the *regular l-recent-cost* of ALG on σ at time t as

$$\mathrm{ALG}^r\big(\sigma; (t - l, t]\big) := \sum_{i:\sigma_i \text{ is a regular request of ALG in } (t-l,t]} p_i. \quad (5)$$

Restrictions on OPT

Without loss of generality, we assume that an online algorithm ALG or an optimal offline algorithm OPT considers purchasing Bahncards at the times of regular requests *only*. The rationale is that the purchase of a Bahncard at any other time can always be delayed to the next regular request without increasing the total cost.

Lemma 1

([1]) At any time t, if $c\big(\sigma; [t, t+T)\big) \geq \gamma := C/(1-\beta)$, OPT has at least one reduced request in $[t, t+T)$.

γ is known as the break-even point.

Corollary 2

At any time t, if the T-future-cost $c\big(\sigma; [t, t+T)\big) < \gamma$, OPT does not purchase a Bahncard at t.

An Optimal Offline Algorithm

An OPT can be obtained by finding the shortest (s, t)-path in a graph constructed with the given travel request sequence σ (can be computed with dynamic programming).

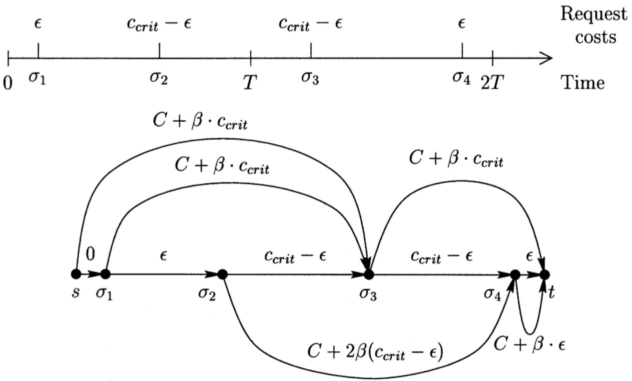


Figure 1: A travel request sequence and the constructed graph (see [1]).

An Optimal Online Algorithm: SUM

An optimal *deterministic* online algorithm for the Bahncard problem is SUM, which is $(2-\beta)$-competitive [1].[4] SUM purchases a Bahncard at a regular request (t, p) whenever its regular T-recent-cost at time t is at least γ, i.e.,

$$\text{SUM}^r\big(\sigma; (t-T, t]\big) \geq \gamma. \tag{6}$$

SUM is $(2-\beta)$-competitive ($\mathtt{CR}_{\text{SUM}} = \frac{C+\gamma}{C+\beta\gamma} = 2-\beta$).

Notably, $2-\beta$ is best competitive ratio a deterministic online algorithm can hope (see [1]).

[4] A randomized online algorithm can have smaller competitive ratio, which will not be covered in this slide.

Outline

How SUM_w Works

SUM purchases a Bahncard based on the cost incurred in the past only. Our first attempt shifts the cost consideration for Bahncard purchasing towards the future with the help of predictions.

SUM_w

At each regular request (t, p), SUM_w predicts the total regular cost in a prediction window $(t, t + w]$ where w $(0 < w < T)$ is the length of the prediction window. SUM_w purchases a Bahncard at a regular request (t, p) whenever the sum of the regular $(T - w)$-recent-cost at t and the predicted total regular cost in $(t, t + w]$ is at least γ, i.e.,

$$\text{SUM}_w^r\big(\sigma; (t + w - T, t]\big) + \hat{c}\big(\sigma; (t, t + w]\big) \geq \gamma. \quad (7)$$

SUM_w reduces to SUM when $w = 0$.

SUM_w is at least $(3-\beta)/(1+\beta)$-Consistent

Unfortunately, the consistency of SUM_w is at least $(3-\beta)/(1+\beta)$, which is even larger than SUM's competitive ratio of $2-\beta$ since $\beta < 1$.

As shown in Fig. 2, consider a travel request sequence σ of five requests: (t_1, ϵ), $(t_2, \gamma-\epsilon)$, $(t_3, \gamma-2\epsilon)$, (t_4, ϵ), and (t_5, ϵ), where $t_1 < t_2 \le t_1 + w < t_4 + w - T < t_1 + T < t_3 < t_4 < t_2 + T < t_3 + w < t_5 \le t_4 + w$.

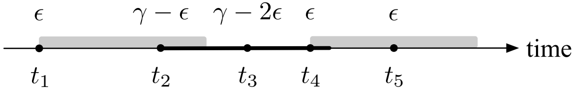


Figure 2: An example where the shaded rectangle (resp. bold line) is the valid time of a Bahncard purchased by SUM_w (resp. OPT).

SUM_w is at least $(3-\beta)/(1+\beta)$-Consistent

For this σ, SUM_w purchases two Bahncards at times t_1 and t_4 respectively, while OPT purchases a Bahncard at time t_2. SUM_w purchases a Bahncard at t_1 because $t_2 \leq t_1 + w$ and $\epsilon + (\gamma - \epsilon) \geq \gamma$. SUM_w does not purchase a Bahncard at t_3 since $t_3 + w < t_5$ and $\text{SUM}^r_w(\sigma; (t_3+w-T, t_3]) + \hat{c}(\sigma; (t_3, t_3+w]) = (\gamma - 2\epsilon) + \epsilon < \gamma$. SUM_w purchases a Bahncard at t_4 because $t_3, t_4, t_5 \in (t_4 + w - T, t_4 + w]$ and the total ticket cost is γ.

When $\epsilon \to 0$, we have

$$\frac{\text{SUM}_w(\sigma)}{\text{OPT}(\sigma)} = \frac{2C + \beta(\gamma + 2\epsilon) + (\gamma - 2\epsilon)}{C + \beta(2\gamma - 2\epsilon) + 2\epsilon} \to \frac{2C + \beta\gamma + \gamma}{C + 2\beta\gamma} = \frac{3-\beta}{1+\beta}. \tag{8}$$

Reviewing SUM_w

SUM_w fails because if a Bahncard is purchased at time t, it is possible that most of the ticket cost in the interval $(t + w - T, t + w]$ is incurred before t (e.g., the Bahncard purchased at time t_4 in Fig. 2). Consequently, only a small fraction of the ticket cost is incurred from t onward and can benefit from the Bahncard purchased. As a result, SUM_w suffers from the same deficiency as SUM.

We remark that it is not helpful to set the prediction window length w to T, because the travel request arising at time $t + T$ (if any) is not covered by the Bahncard purchased at time t.

SUM_w does not have any bounded robustness.

How FSUM Works

What we have learned from SUM_w is that the Bahncard purchase condition should not be based on the total ticket cost in a past time interval and a future prediction window.

Thus, our second algorithm FSUM (Future SUM) is designed to purchase a Bahncard at a regular request (t, p) whenever the predicted T-future-cost at time t is at least γ, i.e.,

$$\hat{c}\big(\sigma; [t, t+T)\big) \geq \gamma. \tag{9}$$

Note that FSUM $\neq \mathrm{SUM}_T$ because the Bahncard purchase condition of SUM_T is $\hat{c}(\sigma; (t, t+T]) \geq \gamma$.

FSUM is $2/(1+\beta)$-Consistent

Theorem 3
FSUM is $2/(1+\beta)$*-consistent.*

Proof.
When both FSUM and OPT do not have a valid Bahncard, they pay the same cost for travel requests. Thus, we focus on analyzing the cost ratio between FSUM and OPT in the time intervals in which at least one of them has a valid Bahncard.

FSUM is $2/(1+\beta)$-Consistent

Theorem 3

FSUM is $2/(1+\beta)$*-consistent.*

Proof.

To do so, we divide the timespan into epochs, denoted as $E_j := [\mu_j, \mu_{j+1})$, where μ_j represents the time when FSUM purchases its j-th Bahncard. Each epoch E_j is further segmented into two phases: the *on* phase $[\mu_j, \mu_j + T)$ where a valid Bahncard is held by FSUM, and the *off* phase $[\mu_j + T, \mu_{j+1})$ where no valid Bahncard is held by FSUM.

In an on phase, OPT purchases at most one Bahncard, while in an off phase, OPT does not make any Bahncard purchase due to the assumption of perfect predictions and Corollary 2.

FSUM is $2/(1+\beta)$-Consistent

Theorem 3

FSUM is $2/(1+\beta)$*-consistent.*

Proof.

If OPT purchases a Bahncard in the on phase of epoch E_j, there are three cases to consider.

Case I. If the Bahncard expires within the off phase of epoch E_j, we analyze the cost ratio in epoch E_j.

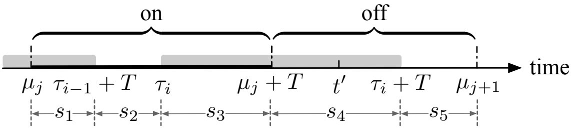


Figure 3: Case I. The shaded rectangle is the valid time of a Bahncard purchased by OPT.

FSUM is $2/(1+\beta)$-Consistent

Theorem 3

FSUM is $2/(1+\beta)$*-consistent.*

Proof.

Case II. If the Bahncard expires during the on phase of the next epoch E_{j+1} and OPT does not purchase any new Bahncard in that on phase, we analyze the cost ratio in epochs E_j and E_{j+1} together.

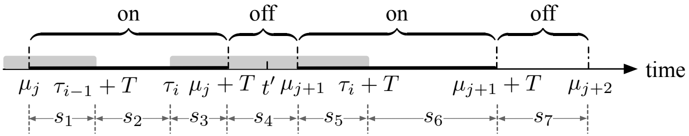


Figure 3: Case II. The shaded rectangle is the valid time of a Bahncard purchased by OPT.

FSUM is $2/(1+\beta)$-Consistent

Theorem 3

FSUM is $2/(1+\beta)$*-consistent.*

Proof.

Case III. If OPT purchases another Bahncard in the on phase of E_{j+1}, we move on to find the first epoch E_{j+x} $(x \geq 1)$ in which Case I or II happens, and then analyze the cost ratio in all epochs $E_j, E_{j+1}, ..., E_{j+x}$ by using the results of Cases I and II.

FSUM is $2/(1+\beta)$-Consistent

Theorem 3

FSUM is $2/(1+\beta)$*-consistent.*

Proof.

If OPT does not purchase a Bahncard in the on phase of E_j, it must have purchased a Bahncard in the on phase of E_{j-1} and the Bahncard expires in the on phase of E_j, due to the assumption of perfect predictions and Lemma 1. Thus, this falls into one of the three cases above.

In all cases, the cost ratio between FSUM and OPT over all these intervals is capped at $2/(1+\beta)$ (e.g., $\frac{C+\beta\gamma+\gamma}{C+2\beta\gamma}$, $\frac{2C+2\beta\gamma}{C+2\beta\gamma}$, etc.), which completes the proof. □

FSUM has no Bounded Robustness

FSUM's consistency generally better than SUM's competitive ratio since $2/(1+\beta) < 2-\beta$ always holds for $0 < \beta < 1$.

However, similar to SUM_w, FSUM does not have any bounded robustness. Consider a scenario where only one travel request (t, p) arises with $p \to 0$, but the predictor yields $\hat{c}(\sigma; [t, t+T)) \geq \gamma$. Then, FSUM purchases a Bahncard at (t, p) and $\mathrm{FSUM}(\sigma)/\mathrm{OPT}(\sigma) = (C + \beta p)/p \to \infty$.

PFSUM Design Considerations

FSUM fails to achieve any bounded robustness because it completely ignores the historical information in the Bahncard purchase condition. Thus, the worst case is that the actual ticket cost in the prediction window is close to 0, while the predictor forecasts that it exceeds γ, in which case hardly anything benefits from the Bahncard purchased.

On the other hand, we note that SUM achieves a decent competitive ratio because a Bahncard is purchased only when the regular T-recent-cost is at least γ, so that the Bahncard cost can be charged to the regular T-recent-cost in the competitive analysis.

How PFSUM Works

Based on the above, we introduce a new algorithm PFSUM (Past and Future SUM), in which the Bahncard purchase condition incorporates the ticket costs in both a past time interval and a future prediction window, but uses them *separately* rather than taking their sum.

PFSUM

PFSUM purchases a Bahncard at a regular request (t, p) whenever (i) the T-recent-cost at t is at least γ, and (ii) the predicted T-future-cost at t is also at least γ:

$$\begin{cases} c(\sigma; (t-T, t]) \geq \gamma, \\ \hat{c}(\sigma; [t, t+T)) \geq \gamma. \end{cases} \tag{10}$$

Note that PFSUM considers the T-recent-cost, but SUM considers only the regular T-recent-cost.

Preliminaries for Analysis

To analyze PFSUM, we again focus on the time intervals in which at least one of PFSUM and OPT has a valid Bahncard, and analyze the cost ratio between PFSUM and OPT in these intervals.

Given a travel request sequence σ, we use $\mu_1 < \cdots < \mu_m$ to denote the times when PFSUM purchases Bahncards. Accordingly, the timespan can be divided into epochs $E_j := [\mu_j, \mu_{j+1})$ for $0 \leq j \leq m$, where we define $\mu_0 = 0$ and $\mu_{m+1} = \infty$.

Each epoch E_j (except E_0) starts with an *on* phase $[\mu_j, \mu_j + T)$ (the valid time of the Bahncard purchased by PFSUM), followed by an *off* phase $[\mu_j + T, \mu_{j+1})$ (in which there is no valid Bahncard by PFSUM). Epoch E_0 has an off phase only.

Prediction Error

We define η as the maximum prediction error among all the predictions used by PFSUM:

$$\eta := \max_{(t,p)} \Big| \hat{c}\big(\sigma; [t, t+T)\big) - c\big(\sigma; [t, t+T)\big) \Big|, \tag{11}$$

where (t, p) is a regular request.

Then, for any travel request (t, p) in an off phase, we have

$$\begin{aligned} c\big(\sigma; [t, t+T)\big) - \eta &\leq \hat{c}\big(\sigma; [t, t+T)\big) \\ &\leq c\big(\sigma; [t, t+T)\big) + \eta. \end{aligned} \tag{12}$$

Preliminaries for Analysis

Lemma 4

The total regular cost in an on phase is at least $\gamma - \eta$.

Proof.

If $c(\sigma; [\mu_j, \mu_j + T)) < \gamma - \eta$ for some j, it follows from (12) that

$\hat{c}(\sigma; [\mu_j, \mu_j + T)) \leq c(\sigma; [\mu_j, \mu_j + T)) + \eta < \gamma - \eta + \eta = \gamma$.

By (10), it means that PFSUM would not purchase a Bahncard at time μ_j, leading to a contradiction. Thus, we must have $c(\sigma; [\mu_j, \mu_j + T)) \geq \gamma - \eta$. □

Preliminaries for Analysis

Lemma 5

In an off phase, the total regular cost in any time interval $[t, t+l)$ of length $l \leq T$ is less than $2\gamma + \eta$.

Proof.

Assume on the contrary that $c(\sigma; [t, t+l)) \geq 2\gamma + \eta$. We take the earliest travel request (t', p) in $[t, t+l)$ such that $c(\sigma; [t, t']) \geq \gamma$. This implies $c(\sigma; [t, t')) < \gamma$. Then, we have $c(\sigma; [t', t+l)) = c(\sigma; [t, t+l)) - c(\sigma; [t, t')) > 2\gamma + \eta - \gamma = \gamma + \eta$. Hence, $c(\sigma; (t' - T, t']) \geq c(\sigma; [t, t']) \geq \gamma$ and $c\big(\sigma; [t', t' + T)\big) \geq c(\sigma; [t', t+l)) > \gamma + \eta$. By (12), the latter further leads to $\hat{c}(\sigma; [t', t' + T)) \geq c(\sigma; [t', t' + T)) - \eta > \gamma$, which means that PFSUM should purchase a Bahncard at time t', contradicting that $t' \in [t, t+l)$ is in the off phase. Hence, $c(\sigma; [t, t+T)) < 2\gamma + \eta$ must hold. □

Preliminaries for Analysis

Lemma 6

Suppose a time interval $[t, t+T)$ overlaps with an off phase. Among the total regular cost in $[t, t+T)$, let s_2, s_3 and s_4 denote those in the preceding on phase, the off phase, and the succeeding on phase respectively (see Fig. 3). If $0 \le \eta \le \gamma$, $s_2 \le \gamma$ and $s_4 \le \gamma$, then the total regular cost in $[t, t+T)$ is no more than $2\gamma + \eta$, i.e., $s_2 + s_3 + s_4 \le 2\gamma + \eta$.[5]

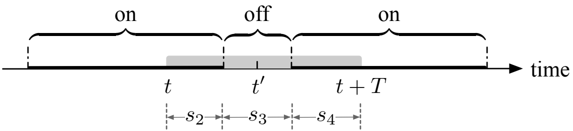


Figure 3: Illustration for Lemma 6. The shaded rectangle is the valid time of a Bahncard purchased by OPT.

[5] The equality can be obtained when $\eta = 0$.

Preliminaries for Analysis

Proof.

Assume on the contrary that $s_2 + s_3 + s_4 > 2\gamma + \eta$. Then, there is at least one travel request during the off phase. We take the earliest travel request (t', p) in the off phase, such that $c(\sigma; [t, t']) \geq \gamma$. With a similar analysis to Lemma 5, we can derive that PFSUM should purchase a Bahncard at time t', contradicting that t' is in the off phase. □

PFSUM Analysis

To analyze PFSUM, we again focus on the time intervals in which at least one of PFSUM and OPT has a valid Bahncard, and analyze the cost ratio between PFSUM and OPT in these intervals.

PFSUM Analysis

Consider a maximal contiguous time interval throughout which at least one of PFSUM and OPT has a valid Bahncard. As shown in Fig. 4, there are 6 different patterns.

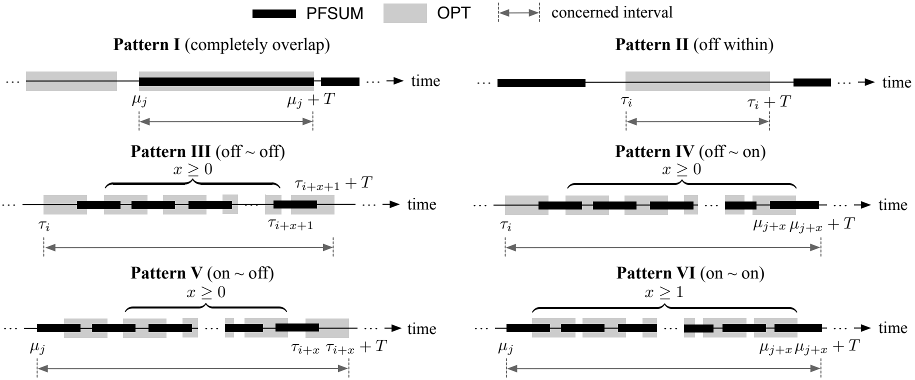


Figure 4: All the 6 patterns of concerned time intervals in which either PFSUM or OPT has a Bahncard.

PFSUM Analysis

If OPT and PFSUM purchase a Bahncard at the same time, the time interval is exactly an on phase (Pattern I), and the cost ratio in it is 1 (Proposition 7).

Otherwise, there are 5 different cases: the time interval does not overlap with any on phase (Pattern II); and the time interval overlaps with at least one on phase – it can start at some time in an on or off phase and end at some time in an on or off phase, giving rise to four cases (Patterns III to VI).

In Patterns II to VI, we assume that none of the involved Bahncards purchased by OPT are bought at the same time as any Bahncards purchased by PFSUM.

Pattern I

Proposition 7

(Pattern I) *If OPT purchases a Bahncard at time* τ_i *at the beginning of epoch* E_j*, i.e.,* $\tau_i = \mu_j$*, then*

$$\frac{PFSUM\big(\sigma; [\mu_j, \mu_j + T)\big)}{OPT\big(\sigma; [\mu_j, \mu_j + T)\big)} = 1. \tag{13}$$

Analyzing Patterns II to VI

If the cost ratios between PFSUM and OPT of Patterns II to VI are all capped by the same bound, the competitive ratio of PFSUM is given by this bound.

Unfortunately, this is *not* exactly true. In what follows, we shall show that the cost ratios of Patterns II to V can be capped by the same bound (Propositions 8 to 11), which is the competitive ratio of PFSUM that we would like to prove.

For Pattern VI, we show that its cost ratio is capped by the same bound *if a particular condition holds*, where we refer to such Pattern VI as *augmented* Pattern VI (Proposition 12).

Analyzing Patterns II to VI

For non-augmented Pattern VI, we show that it must be accompanied by Patterns I to IV in the sense that a time interval of non-augmented Pattern VI must be preceded (not necessarily immediately) by a time interval of Pattern I, II, III or IV.

We prove that the cost ratio of non-augmented Pattern VI combined with Pattern I, II, III or IV is capped by the same aforesaid bound (Propositions 13 to 14). This then completes the analysis and shows that the competitive ratio of PFSUM is given by this bound.

Pattern II

Proposition 8

(Pattern II) *If OPT purchases a Bahncard at time* τ_i *in the off phase of an epoch* E_j *and the Bahncard expires in the same off phase, i.e.,* $\mu_j + T \leq \tau_i < \tau_i + T < \mu_{j+1}$*, then*

$$\frac{PFSUM\big(\sigma; [\tau_i, \tau_i + T)\big)}{OPT\big(\sigma; [\tau_i, \tau_i + T)\big)} < \frac{2\gamma + \eta}{(1+\beta)\gamma + \beta\eta}. \qquad (14)$$

Proof.

By Lemma 5, $c(\sigma; [\tau_i, \tau_i + T)) < 2\gamma + \eta$. The result is then obtained immediately. □

Analyzing Patterns III to VI

The proof technique of the following propositions for Patterns III to VI is to divide the time interval concerned into sub-intervals, where each sub-interval starts and ends at the time when OPT or PFSUM purchases a Bahncard or a Bahncard purchased by OPT or PFSUM expires.

Then, for $0 \leq \eta \leq \gamma$ and $\eta > \gamma$, we respectively derive the upper bound of the cost ratio based on Lemmas 4 to 6. All the bounds in Propositions 9 to 12 are tight (achievable).

Pattern III

Proposition 9

(Pattern III) *Denote by* τ_i *and* τ_{i+x+1} *respectively the first and last Bahncards purchased by OPT in Pattern III. The cost ratio in the time interval of Pattern III satisfies*

$$\frac{PFSUM(\sigma;[\tau_i,\tau_{i+x+1}+T))}{OPT(\sigma;[\tau_i,\tau_{i+x+1}+T))} \leq \begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{15}$$

Pattern IV

Proposition 10

(Pattern IV) *Denote by* τ_i *and* μ_{j+x} *respectively the first Bahncard purchased by OPT and the last Bahncard purchased by PFSUM in Pattern IV. The cost ratio in the time interval of Pattern IV satisfies*

$$\frac{PFSUM\big(\sigma;[\tau_i,\mu_{j+x}+T)\big)}{OPT\big(\sigma;[\tau_i,\mu_{j+x}+T)\big)} \leq \begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{16}$$

Pattern V

Proposition 11

(Pattern V) *Denote by μ_j and τ_{i+x} respectively the first Bahncard purchased by PFSUM and the last Bahncard purchased by OPT in Pattern V. The cost ratio in the time interval of Pattern V satisfies*

$$\frac{PFSUM(\sigma;[\mu_j,\tau_{i+x}+T))}{OPT(\sigma;[\mu_j,\tau_{i+x}+T))} \leq \begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{17}$$

Augmented Pattern VI

We refer to Pattern VI as augmented Pattern VI if the total regular cost in the last on phase involved is at least γ. Proposition 12 shows that the cost ratio of augmented Pattern VI is capped by the same bound as Patterns III to V.

Proposition 12

(Augmented Pattern VI) *Denote by μ_j and μ_{j+x} respectively the first and last Bahncards purchased by PFSUM in Pattern VI. If the total regular cost in the on phase of E_{j+x} is at least γ, the cost ratio in the time interval of Pattern VI satisfies*

$$\frac{PFSUM(\sigma;[\mu_j,\mu_{j+x}+T))}{OPT(\sigma;[\mu_j,\mu_{j+x}+T))} \leq \begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{18}$$

Non-Augmented Pattern VI

Now, we examine non-augmented Pattern VI. Note that the time intervals of Patterns V and VI cannot exist in isolation. By the definition of PFSUM, when a Bahncard is purchased at time μ_j, the total regular cost in the preceding interval $(\mu_j - T, \mu_j]$ is at least γ. Consequently, by Lemma 1,[6] OPT must purchase a Bahncard whose valid time overlaps with $(\mu_j - T, \mu_j]$.

To deal with non-augmented Pattern VI that starts at time μ_j for some j, we backtrack from μ_j to find out what happens earlier. Our target is to identify all possible patterns that might precede Pattern VI.

[6] It is obvious that Lemma 1 is also applicable to a left-open and right-closed interval of length T.

Pattern Graph

Note that the time interval $(\mu_j - T, \mu_j]$ definitely intersects with the off phase of epoch E_{j-1} and may also intersect with the on phase of E_{j-1} (if the off phase of E_{j-1} is shorter than T). Therefore, it is possible for all Patterns I to VI to precede Pattern VI.

If Pattern I, IV or VI precedes Pattern VI, there is an off phase in between, in which neither PFSUM nor OPT holds a valid Bahncard.

Pattern Graph

Fig. 5 presents a *pattern graph* to illustrate all possible concatenations of patterns preceding Pattern VI. In this graph, a node represents a pattern, and an edge from node *i* to node *j* means pattern *i* can precede pattern *j*.

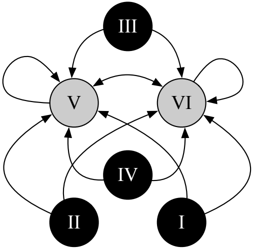


Figure 5: Pattern graph.

Since every Pattern V or VI must be preceded by some pattern, the backtracking will always encounter a time interval of Pattern I, II, III or IV. We stop backtracking at the first Pattern I, II, III or IV encountered.

Backtracking

We use $p_1 \oplus p_2$ to denote the composite of pattern p_1 followed by pattern p_2; use p^y to denote a sequence comprising y consecutive instances of pattern p; and use $\{p_1, \ldots, p_n\} \oplus p_j$ to represent all possible composite patterns of the form $p_i \oplus p_j$ for each $i = 1, \ldots, n$.

Then, the patterns encountered in the backtracking can be represented by

$$\{\mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}\} \oplus \{\mathrm{V}, \mathrm{VI}\}^y \oplus \mathrm{VI}, \tag{19}$$

where $y \geq 0$.

Backtracking

- Following the earlier argument, for each VI in $\{\text{V}, \text{VI}\}^y$ of (19), the total regular cost in the last on phase of Pattern VI and the following off phase is at least γ. It is easy to see that the cost ratio between PFSUM and OPT in the time interval of such Pattern VI is no larger than that of augmented Pattern VI and hence the upper bound given in Proposition 12.
- For each V in $\{\text{V}, \text{VI}\}^y$ of (19), the cost ratio of Pattern V is capped by the same bound based on Proposition 11.

Now, we only need to prove that the cost ratio of non-augmented Pattern VI combined with Pattern I, II, III or IV at the beginning of (19) is also capped by the same bound.

Pattern VI Proceeded by Pattern II or III

Proposition 13

The cost ratio in the combination of a time interval of Pattern VI and a time interval of Pattern II or III is bounded by

$$\begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases}$$

Pattern VI Proceeded by Pattern I or IV

Proposition 14

The cost ratio in the combination of a time interval of Pattern VI and a time interval of Pattern I or IV encountered in backtracking is bounded by

$$\begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{20}$$

Proof.

Consider Pattern IV. When analyzing the cost ratio of Pattern IV in Proposition 10, it is assumed that the total regular cost in the last on phase is at least $\gamma - \eta$. In IV of (19), the total cost of travel requests in the last on phase of Pattern IV and the following off phase is at least γ, following the earlier argument. Thus, there is an additional cost of at least η.

Pattern VI Proceeded by Pattern I or IV

Proposition 14

The cost ratio in the combination of a time interval of Pattern VI and a time interval of Pattern I or IV encountered in backtracking is bounded by

$$\begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \le \eta \le \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{20}$$

Proof.

Therefore, we can "migrate" an additional cost of η to the non-augmented Pattern VI at the end of (19) and make the latter become an augmented Pattern VI so that the result of Proposition 12 can be applied. Meanwhile, the proof of Proposition 10 still applies to the Pattern IV even if the cost η is removed from the last on phase.

Pattern VI Proceeded by Pattern I or IV

Proposition 14

The cost ratio in the combination of a time interval of Pattern VI and a time interval of Pattern I or IV encountered in backtracking is bounded by

$$\begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \leq \eta \leq \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{20}$$

Proof.

Hence, after migrating the cost of η, the Patterns IV and VI involved have the same upper bound of cost ratio given in (20). □

Competitive Ratio of PFSUM

By the above analysis and Propositions 8 to 14, we have:

Theorem 15

PFSUM has a competitive ratio of

$$\mathcal{CR}_{PFSUM}(\eta) = \begin{cases} \frac{2\gamma+(2-\beta)\eta}{(1+\beta)\gamma+\beta\eta} & 0 \le \eta \le \gamma, \\ \frac{(3-\beta)\gamma+\eta}{(1+\beta)\gamma+\beta\eta} & \eta > \gamma. \end{cases} \tag{21}$$

By letting $\eta = 0$, it is easy to see that the consistency of PFSUM is $2/(1+\beta)$. Note that the competitive ratio of (21) is a continuous function of the prediction error η. It increases from $2/(1+\beta)$ to $(4-\beta)/(1+2\beta)$ as η increases from 0 to γ, and further increases from $(4-\beta)/(1+2\beta)$ towards $1/\beta$ as η increases from γ towards infinity. Hence, PFSUM is $1/\beta$-robust.

Outline

Experimental Settings

We conduct extensive experiments comparing PFSUM with SUM [1], FSUM, and PDLA [15], under various parameter settings.

To accommodate PDLA, we discretize time while setting a sufficiently large timespan, closely approximating a continuous time scenario.

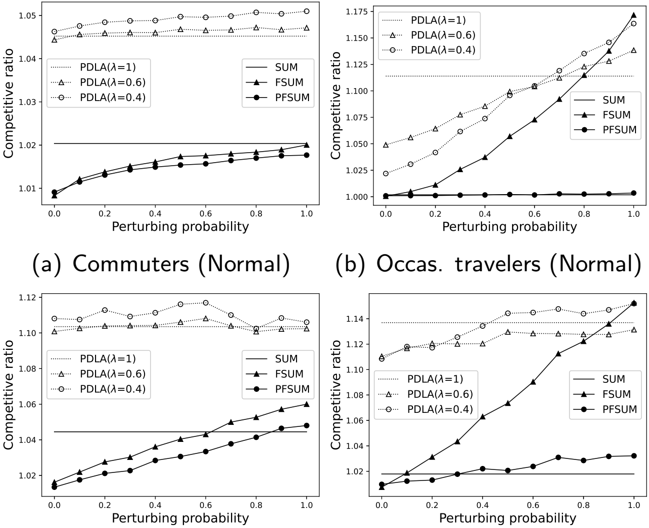
- **Input distributions.**
  - Referring to the experimental setup of [17], we consider two main types of traveler profiles: *commuters*, and *occasional travelers*. We model the inter-request time of occasional travelers using an exponential distribution with a mean of 2, i.e. $\exp(\lambda = \frac{1}{2})$.
  - For each generated request sequence, we investigate three types of ticket price distributions: a Normal distribution, a Uniform distribution, and a Pareto distribution.

Experimental Settings

- **Noisy prediction.**
  - The predictions are generated by adding noise to the original instance, following the methodology used by [15].
  - For each day in a given instance, there is a probability p of removing the travel request on that day, if it exists.
  - Meanwhile, there is also a probability p of adding random noise, sampled from the same distribution used for generating ticket prices, to the price of the travel request, or simply adding a travel request if no travel request exists on that day.

These two operations are executed independently. The total regular cost of the perturbed instance in the interval $[t, t+T)$ is then used as the prediction $\hat{c}(\sigma; [t, t+T))$ at time (day) t. Intuitively, the prediction error increases with the perturbation probability.

Experimental Results



(a) Commuters (Normal) (b) Occas. travelers (Normal)

(c) Commuters (Pareto) (d) Occas. travelers (Pareto)

Figure 6: Experimental results.

References I

[1] Rudolf Fleischer. "On the Bahncard problem". In: *Theoretical Computer Science* 268.1 (2001), pp. 161–174.

[2] Michael Mitzenmacher and Sergei Vassilvitskii. "Algorithms with predictions". In: *Communications of the ACM* 65.7 (2022), pp. 33–35.

[3] Manish Purohit, Zoya Svitkina, and Ravi Kumar. "Improving online algorithms via ML predictions". In: *Advances in Neural Information Processing Systems* 31 (2018).

[4] Sreenivas Gollapudi and Debmalya Panigrahi. "Online Algorithms for Rent-Or-Buy with Expert Advice". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, 2019, pp. 2319–2327.

References II

[5] Alexander Wei and Fred Zhang. "Optimal Robustness-Consistency Trade-offs for Learning-Augmented Online Algorithms". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., 2020, pp. 8042–8053.

[6] Shufan Wang, Jian Li, and Shiqiang Wang. "Online Algorithms for Multi-shop Ski Rental with Machine Learned Advice". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., 2020, pp. 8150–8160.

References III

[7] Keerti Anand, Rong Ge, and Debmalya Panigrahi. "Customizing ML Predictions for Online Algorithms". In: *Proceedings of the 37th International Conference on Machine Learning*. Ed. by Hal Daumé III and Aarti Singh. Vol. 119. Proceedings of Machine Learning Research. PMLR, 2020, pp. 303–313.

[8] Sungjin Im et al. "Non-clairvoyant scheduling with predictions". In: *Proceedings of the 33rd ACM Symposium on Parallelism in Algorithms and Architectures*. 2021, pp. 285–294.

References IV

[9] Yossi Azar, Stefano Leonardi, and Noam Touitou. "Flow Time Scheduling with Uncertain Processing Time". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, 1070–1080. ISBN: 9781450380539. DOI: 10.1145/3406325.3451023.

[10] Honghao Lin, Tian Luo, and David Woodruff. "Learning Augmented Binary Search Trees". In: *Proceedings of the 39th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri et al. Vol. 162. Proceedings of Machine Learning Research. PMLR, 2022, pp. 13431–13440.

[11] Mozhengfu Liu and Xueyan Tang. "Dynamic Bin Packing with Predictions". In: *Proc. ACM Meas. Anal. Comput. Syst.* 6.3 (2022). DOI: 10.1145/3570605.

References V

[12] Justin Chen et al. "Faster Fundamental Graph Algorithms via Learned Predictions". In: *Proceedings of the 39th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri et al. Vol. 162. Proceedings of Machine Learning Research. PMLR, 2022, pp. 3583–3602.

[13] Sungjin Im et al. "Online Dynamic Acknowledgement with Learned Predictions". In: *arXiv preprint arXiv:2305.18227* (2023).

[14] Thodoris Lykouris and Sergei Vassilvtiskii. "Competitive Caching with Machine Learned Advice". In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. 2018.

References VI

[15] Etienne Bamas, Andreas Maggiori, and Ola Svensson. “The Primal-Dual method for Learning Augmented Algorithms”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., 2020, pp. 20083–20094.

[16] Niv Buchbinder, Kamal Jain, and Joseph Naor. “Online primal-dual algorithms for maximizing ad-auctions revenue”. In: *European Symposium on Algorithms*. Springer. 2007, pp. 253–264.

[17] Mike Timm and Sabine Storandt. “On the multi-kind BahnCard problem”. In: (2020).