Supplementary Materials for Online Stochastic Generalized Assignment Problem with Demand Learning

Yanwen Li Zihao Li Limeng Liu Hao Wang Zhenzhen Yan
Nanyang Technological University
{yanwen.li, hao_wang, yanzz}@ntu.edu.sg
{zihao004, limeng001}@e.ntu.edu.sg

A Additional Lemmas

In our analysis in Section 4, we need the following two lemmas to facilitate the estimation of the actual demand distribution using the empirical distribution and the number of arrivals of each online item type. The second lemma is just a corollary of Chernoff bound.

Lemma A.1 (Estimation of Distribution, [Dvoretzky et al., 1956]). *Suppose $X_1, \ldots, X_t$ are independent and identically distributed random variables with cumulative distribution function $F(\cdot)$. Denote F_t as the associated empirical distribution, where $F_t(x) = \frac{1}{t}\sum_{i=1}^{t} \boldsymbol{1}_{\{X_i \leq x\}}$, for all real number x. Here, $\boldsymbol{1}_{\{X_i \leq x\}}$ is the indicator function which takes value 1 if $X_i \leq x$, otherwise 0. For any $\epsilon > 0$, $Pr[\sup_{x\in\mathbb{R}} |F_t(x) - F(x)| > \epsilon] \leq 2e^{-2t\epsilon^2}$.*

Lemma A.2 (Number of Samples, [Mitzenmacher and Upfal, 2017]). *Suppose $X_1, \ldots, X_t$ are independent and identically distributed random variables taking values in $\{0, 1\}$. Denote X as their sum and μ as the expected value of the sum. Then, $Pr[X \geq \frac{1}{2}\mu] \geq 1 - e^{-\frac{\mu}{8}}$.*

B Missing Proofs in Section 2

B.1 Proof of Lemma 2.1

Proof. Since Theorem 1 in Alaei et al. [2013] is a generalized version of the model we consider in this LP, where we only allow the i.i.d. arrival distributions and the demand distributions of one fixed type for different online arrivals are the same. It suffices to show we can build one solution with the same objective value in our LP as that in the LP ($\overline{OPT}$) mentioned in Theorem 1 in Alaei et al. [2013]. We assume the solution of LP $\overline{OPT}$ is $\{x^*_{t,v,u}\}$, where $t \in [T], v \in V$ and $u \in U$. Here, we modify their notations correspondingly based on our notations. From their first constraints $\sum_{t\in[T]}\sum_{v\in V} \tilde{s}_{t,v,u} x^*_{t,v,u} \leq c_u$ for each $u \in U$, under our model, we have $\sum_{v\in V} d_v (\sum_{t\in[T]} x^*_{t,v,u}) \leq c_u$ because all $\tilde{s}_{t,v,u}$s in their model are exactly d_v in our model.

For their second constraints $\sum_{u\in U} x^*_{t,v,u} \leq p_{t,v}$ for each $t \in [T]$ and $v \in V$, under our model, we can sum up for all $t \in [T]$ and get $\sum_{u\in U}(\sum_{t\in[T]} x^*_{t,v,u}) \leq p_v T$ for each $v \in V$. This is from the assumption of i.i.d. arrivals.

Thus, if we choose $x_{u,v}$ in our model as $\sum_{t\in[T]} x^*_{t,v,u}$, we find this satisfies all our constraints, and the objective $\sum_{u\in U, v\in V} r_{u,v} x_{u,v} = \sum_{u\in U, v\in V, t\in[T]} r_{u,v} x^*_{t,v,u}$, where the latter is the same objective in their LP by modifying the notations. This finishes the proof. □

Preprint. Under review.

C Missing Proofs in Section 3

C.1 Proof of Theorem 3.4

Before we start our proof, we first simplify some notations based on Bernoulli distributions. Specifically, we assume 0 appears with the probability p and $\tilde{p}$ in the demand distributions D and $\tilde{D}$ respectively. We have $p \geq \tilde{p}$ based on the definitions.

Proof. We first prove that $F_{\tilde{W}_t}(x) \leq F_{W_t}(x)$ for all x and $t \in [T]$. From the Bernoulli distribution, we only consider the integer x without loss of generality. We prove this by induction. For $t = 1$, this is satisfied since both $\tilde{W}_1$ and W_t can take the value 0. For $t+1$, if $x < \theta_t$, we have $F_{\tilde{W}_{t+1}}(x) = F_{\tilde{W}_t}(x-1) + (F_{\tilde{W}_t}(x) - F_{\tilde{W}_t}(x-1))\tilde{p} \leq F_{\tilde{W}_t}(x-1) + (F_{\tilde{W}_t}(x) - F_{\tilde{W}_t}(x-1))p$ and $F_{W_{t+1}}(x) = F_{W_t}(x-1) + (F_{W_t}(x) - F_{W_t}(x-1))p$, from the modified γ-magician procedure. Since $F_{\tilde{W}_t}(x-1) \leq F_{W_t}(x-1)$ and $F_{\tilde{W}_t}(x) \leq F_{W_t}(x)$ and $p \in [0,1]$, we can conclude that $F_{\tilde{W}_{t+1}}(x) \leq F_{W_{t+1}}(x)$.

If $x = \theta_t$, by the procedure, we have $F_{\tilde{W}_{t+1}}(x) = F_{\tilde{W}_t}(x-1) + (F_{\tilde{W}_t}(x) - F_{\tilde{W}_t}(x-1))(\tilde{p}r_t + (1-r_t)) \leq F_{\tilde{W}_t}(x-1) + (F_{\tilde{W}_t}(x) - F_{\tilde{W}_t}(x-1))(pr_t + (1-r_t))$ and $F_{W_{t+1}}(x) = F_{W_t}(x-1) + (F_{W_t}(x) - F_{W_t}(x-1))(pr_t + (1-r_t))$. Since $pr_t + (1-r_t) \in [0,1]$, with the same analysis, we can also get $F_{\tilde{W}_{t+1}}(x) \leq F_{W_{t+1}}(x)$.

For the last case where $x > \theta_t$, from the monotonicity of the choice of θ_t, we directly have $F_{\tilde{W}_{t+1}}(x) = 1 = F_{W_{t+1}}(x)$.

We then can check whether $\Pr[\tilde{I}_t] \leq \Pr[I_t]$ holds for all $t \in [T]$, which finishes the proof according to Claim 3.3. Under the condition that $\gamma \leq 1 - \frac{1}{\sqrt{k+3}}$, from Claim 3.2 and Claim 3.3, we have $\Pr[\tilde{I}_t] = \gamma = \Pr[\tilde{W}_t < \theta_t] + r_t \Pr[\tilde{W}_t = \theta_t] = F_{\tilde{W}_t}(\theta_t - 1) + r_t(F_{\tilde{W}_t}(\theta_t) - F_{\tilde{W}_t}(\theta_t - 1))$ and $\Pr[I_t] = F_{W_t}(\theta_t - 1) + r_t(F_{W_t}(\theta_t) - F_{W_t}(\theta_t - 1))$. From the same analysis, we get $\Pr[I_t] \geq \Pr[\tilde{I}_t] = \gamma$ holds for all $t \in [T]$. □

D Missing Proofs in Section 4

D.1 Proof of Lemma 4.1

Proof. From Lemma A.2, for each type v, we can treat each random variable X_i taking value 1 with a probability of p_v, where $i \in [T_0]$. We have the number of arrivals of items of type v is at least $\frac{T_0 p_v}{2} \geq \frac{\alpha N}{2}$ with a probability of at least $1 - e^{-\frac{T_0 p_v}{8}}$. By union bound, we have: with a probability of at least $1 - \sum_{v \in V} e^{-\frac{T_0 p_v}{8}} \geq 1 - me^{-\frac{\alpha N}{8}}$, the number of arrivals of items of each type is at least $\frac{\alpha N}{2}$. □

D.2 Proof of Lemma 4.2

Proof. Conditioning on E where the number of arrivals of items of each type in the sampling phase is at least $\frac{\alpha N}{2}$, from Lemma 2.2, for a fixed $v \in V$ and a $\delta > 0$, considering each X_i is the random variable corresponding to the realized demand of one arrival of type v during the sampling phase, we have $\bar{d}_v \leq (1+\delta)d_v$ with a probability of at least $1 - e^{-2\delta^2 d_v^2 \frac{\alpha N}{2}} \geq 1 - e^{-\delta^2 \underline{d}^2(\alpha N)}$. Since we get $\tilde{d}_v$ from the sum of ϵ and $\bar{d}_v$, we have $\tilde{d}_v \leq (1+\delta)d_v + \epsilon \leq (1 + \delta + \epsilon/\underline{d})d_v$ with probability $1 - e^{-\delta^2 \underline{d}^2(\alpha N)}$. By union bound over all $v \in V$, we finish our proof. □

D.3 Proof of Lemma 4.3

Proof. Conditioning on E where the number of arrivals of items of each type in the sampling phase is at least $\frac{\alpha N}{2}$, from Lemma A.1, for a fixed $v \in V$, we treat each X_i as the corresponding random variable of the realized demand of each arrival of type v in the sampling phase, and we have $|\bar{F}_v(x) - F_v(x)| \leq \epsilon$ holds for all x with a probability of $1 - 2e^{-2\frac{\alpha N}{2}\epsilon^2} = 1 - 2e^{-\epsilon^2(\alpha N)}$. Here, we denote $\bar{F}_v(x)$ as the empirical distribution calculated at Step 15 of our algorithm. From the modification of $\bar{F}_v(x)$ at Step 16 of our algorithm, we have $\tilde{F}_v(x) \leq F_v(x)$ holds for all x with probability $1 - 2e^{-\epsilon^2(\alpha N)}$. By union bound, we get the conclusion. □

D.4 Proof of Lemma 4.4

Proof. We assume the optimal solution of $LP(\mathbf{d}, T)$ is $\{x^*_{uv}\}$, where $\mathbf{d}$ represents the expected values of actual demand distributions and T is the total horizon. From Lemma 2.1, we have $\sum_{u \in U, v \in V} r_{uv} x^*_{uv} \geq \text{OPT}$, so it suffices to show $\{\frac{1-\alpha}{1+\delta+\epsilon/\underline{d}} x^*_{uv}\}$ is a solution of $LP(\tilde{\boldsymbol{d}}, T')$.

For Constraints (1a), for each $v \in V$, $\sum_{u \in U} \frac{1-\alpha}{1+\delta+\epsilon/\underline{d}} x^*_{uv} \leq (1-\alpha) \sum_{u \in U} x^*_{uv} \leq p_v(1-\alpha)T = p_v T'$. The first inequality is from $1 + \delta + \epsilon/\underline{d} \geq 1$ and the second inequality is from the corresponding Constraint (1a) of $LP(\mathbf{d}, T)$.

For Constraints (1b), for each $u \in U$, $\sum_{v \in V} \tilde{d}_v \frac{1-\alpha}{1+\delta+\epsilon/\underline{d}} x^*_{uv} \leq (1-\alpha) \sum_{v \in V} d_v x^*_{uv} \leq c_u$. The first inequality is from Lemma 4.2 while the second inequality is from the corresponding Constraint (1b) of $LP(\mathbf{d}, T)$ and the fact that $1 - \alpha \leq 1$. □

D.5 Proof of Lemma 4.5

To prove this lemma, we need to verify that the two conditions defined in the modified GMP are satisfied. By utilizing the matching probability specified in Step 24 and the definition of $F_{\tilde{D}}$, we can verify the first condition $F_{\tilde{D}}(x) \leq F_D(x)$ according to Lemma 4.3. After proving that $\tilde{d}_v$ is an upper bound of the expectation of the distribution whose CDF is $\tilde{F}_v$, we can also verify the second assumption $T \cdot \mathbb{E}[X] \leq k$ according to the definition of $\tilde{D}$ and Constraint (1b).

Proof. We first check the condition $F_{\tilde{D}}(x) \leq F_D(x)$ for each bin $u \in U$. We fix the bin u. From Step 19 of our algorithm, we have $F_{\tilde{D}}(x) = (1 - \sum_{v \in V} \frac{\tilde{x}_{uv}}{T'}) + \sum_{v \in V} \frac{\tilde{x}_{uv}}{T'} \tilde{F}_v(x)$ for each $x \in [0, 1]$. We then consider the calculation of F_D. From Step 24 of our algorithm, for each $x \in [0, 1]$, we have $F_D(x) = (1 - \sum_{v \in V} p_v \frac{\tilde{x}_{uv}}{p_v T'}) + \sum_{v \in V} p_v \frac{\tilde{x}_{uv}}{p_v T'} F_v(x)$. The first term corresponds to the case that this bin u is not chosen and the second term corresponds to the case that the arriving item is of type v and chooses the bin u. Then, by simplification and Lemma 4.3, we get $F_{\tilde{D}}(x) \leq F_D(x)$ for bin u.

We then check the second condition that $T \cdot \mathbb{E}[X] \leq k$. From the definition of the modified GMP, X follows $\tilde{D}$, the T in the definition is the T' here, and k is c_u for the GMP of each bin u. By Steps 10 and 15 of our algorithm, for each $v \in V$, we have $\bar{d}_v = \mathbb{E}[X_v]$, where X_v follows the empirical CDF distribution F_v. Since F_v takes value in $[0, 1]$, after the modification at Step 16 of our algorithm, the expected value can increase at most $(1 - 0) \cdot \epsilon$. With the definition of $\tilde{d}_v$ at Step 11, we have $\tilde{d}_v \geq \mathbb{E}[\tilde{X}_v]$, where $\tilde{X}_v$ follows the CDF $\tilde{F}_v(x)$ defined at Step 16. We now fix a bin u. We then use the definition of the CDF distribution $F_{\tilde{D}}$ at Step 19, then we have $T' \cdot \mathbb{E}[X] = T' \sum_{v \in V} \frac{\tilde{x}_{uv}}{T'} \mathbb{E}[\tilde{X}_v] \leq T' \sum_{v \in V} \frac{\tilde{x}_{uv}}{T'} \tilde{d}_v \leq c_u$. The first equation is from the definition at Step 19 and the last inequality is from Constraint (1b) of $LP(\tilde{\boldsymbol{d}}, T')$. □

D.6 Proof of Theorem 4.6

Proof. From Lemmas 4.1, 4.2, 4.3, and 4.5, by union bound, we have: with a probability of at least $1 - me^{-\frac{\alpha N}{8}} - me^{-\delta^2 \underline{d}^2(\alpha N)} - 2me^{-\epsilon^2(\alpha N)}$, the corresponding problem of each bin u can reduce to a modified GMP problem. If all demand distributions are Bernoulli distributions, by the definitions of $\tilde{D}$ and D, we get these two distributions are also Bernoulli. From Theorem 3.4, we can achieve a competitive ratio of $1 - \frac{1}{\sqrt{c_u+3}} \geq 1 - \frac{1}{\sqrt{k+3}}$ when choosing $\gamma = 1 - \frac{1}{\sqrt{c_u+3}}$ for each bin u, compared to the benchmark OPT' of $LP(\tilde{\boldsymbol{d}}, T')$.

With Lemma 4.4, we get the exact competitive ratio is $\frac{1-\alpha}{1+\delta+\epsilon/\underline{d}}(1 - \frac{1}{\sqrt{k+3}})$. □

E Details in Section 5

E.1 Demand Distributions

Let X denote the random variable that follows a demand distribution D.
(1) D is a Bernoulli distribution $\mathcal{B}(q)$. $\Pr[X = 1] = q$ and $\Pr[X = 0] = 1 - q$. For each online item type, we generate q from a uniform distribution $\mathcal{U}[0, 1]$.

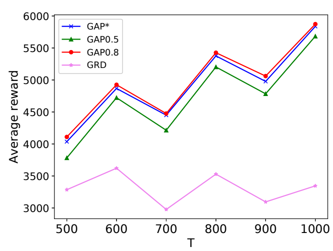
(2) D is a uniform distribution $\mathcal{U}[a, b]$. For each online item type, we generate a and b from a uniform distribution $\mathcal{U}[0, 1]$. The smaller one is used as a, and the bigger one is b.

(3) D is a truncated normal distribution $\mathcal{T}(a, b, \mu, \sigma^2)$. For each online item type, the truncated interval $[a, b]$ is always equal to $[0, 1]$. We generate μ from a uniform distribution $\mathcal{U}[0, 1]$, and σ from a uniform distribution $\mathcal{U}[0, 0.5]$.

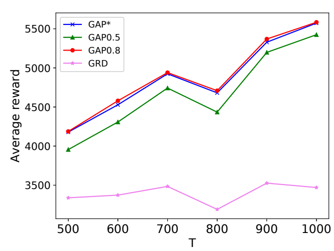
E.2 Supplementary Results

Main results have been discussed in the main paper. In this section, we provide some supplementary experiment results that are omitted in the main paper.

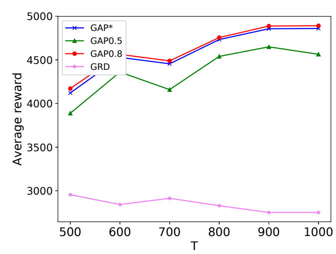
Reward curves We show the average reward of our tested algorithms in Figure E.1 and E.2. In Figure E.1, we fix the initial capacity of each offline bin $k = 10$, and vary the number of online arrivals T from 500 to 1000. We can see the GRD is little affected by T, and our heuristics get more rewards when T goes larger. In Figure E.2, we fix $T = 500$ and test for different $k = 2, 3, \cdots, 10$. We can see that larger k will make all tested algorithms better-performing because the capacity becomes larger, then each bin can serve more items. The gap between every heuristic and GRD becomes larger when k is larger, especially GAP0.5 and GAP* .



(a) Bernoulli distribution

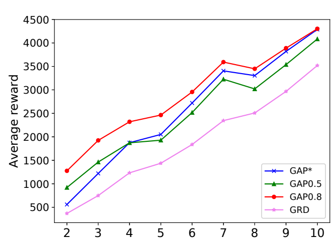


(b) Uniform distribution

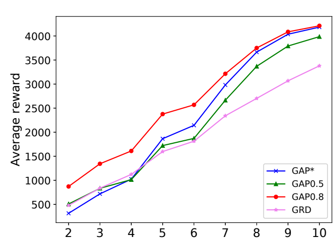


(c) Truncated normal distribution

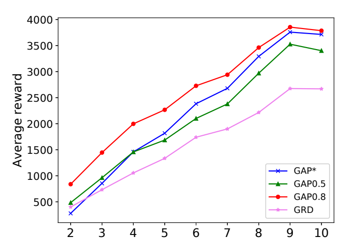
Figure E.1: $k = 10, T = 500, 600, \cdots, 1000$.



(a) Bernoulli distribution



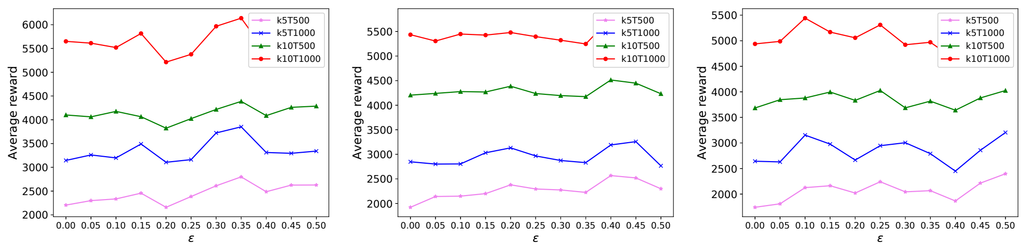
(b) Uniform distribution



(c) Truncated normal distribution

Figure E.2: $T = 500$, $k = 2, 3, \cdots, 10$.

Different ϵ's We test the effect of error parameter ϵ for different (k, T) pairs: $(k = 5, T = 500)$, $(k = 5, T = 1000)$, $(k = 10, T = 500)$ and $(k = 10, T = 1000)$. The results are showed in Figure E.3. We can see that for different ϵ's, the average rewards do not change a lot even under different (k, T)'s and different distributions. When fixing the type of distribution, the shapes of the reward curves of different (k, T)'s are similar, and the best ϵ^* choice is the same. The best ϵ^*'s are different for different distributions, i.e., 0.35 for Bernoulli and 0.10 for truncated normal distribution. We use $\epsilon = 0$ in our main paper because (1) the best ϵ^* for different distributions do not exist; (2) though $\epsilon = 0$ is not the best among all choices, the performance of $\epsilon = 0$ is good enough.



(a) Bernoulli distribution (b) Uniform distribution (c) Truncated normal distribution

Figure E.3: $\epsilon = 0, 0.05, 0.10, \cdots, 0.50$.

Runtime We summarize the runtime of our experiments in Table 1. This table records the average runtime for a single realization of each instance. We can see that the runtime of different heuristics with different distributions are very close and still acceptable compared to that of GRD .

Table 1: Runtime of different algorithms with different distributions

	Runtime (seconds)		
Algorithm	Bernoulli	Uniform	Truncated normal
GRD	0.0016	0.0019	0.0016
GAP*	4.2172	4.0496	3.9455
GAP0.5	4.2409	3.9591	3.8964
GAP0.8	4.2390	4.0444	3.9321

References

Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, pages 642–669, 1956.

Michael Mitzenmacher and Eli Upfal. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, 2017.

Saeed Alaei, MohammadTaghi Hajiaghayi, and Vahid Liaghat. The online stochastic generalized assignment problem. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques: 16th International Workshop, APPROX 2013, and 17th International Workshop, RANDOM 2013, Berkeley, CA, USA, August 21-23, 2013. Proceedings*, pages 11–25. Springer, 2013.