Online Bipartite Matching with budget Refill

M.Cherifa, C.Calauzènes, V.Perchet

December 15, 2023

Motivation: user allocation

Matching on bipartite graphs

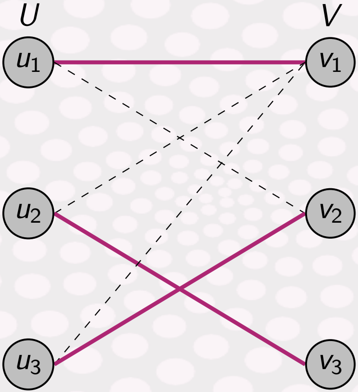
Matching on a Bipartite Graph

Let $G = (U, V, E)$ be a a bipartite graph:

- U and V two sets of vertices.
- Each node in U has a budget $b_u = 1$.
- Edges are **only between** U and V, $E = \{(u, v), u \in U, v \in V\}$.

Matching on a Bipartite Graph

A matching is a subset of E with no common vertices.



Online matching problem

For $t = 1, \ldots, |V|$:

- v_t arrives with its edges.
- the algorithm can match it to a free vertex in U.
- the matching decision is irrevocable.

Evaluating the performance of an Algorithm

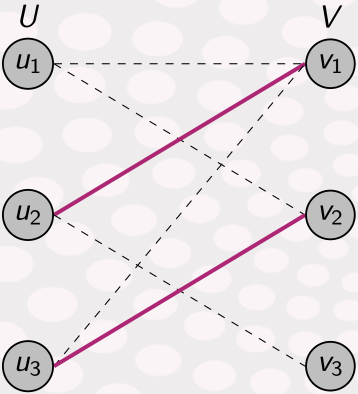


Figure: $ALG = 2$

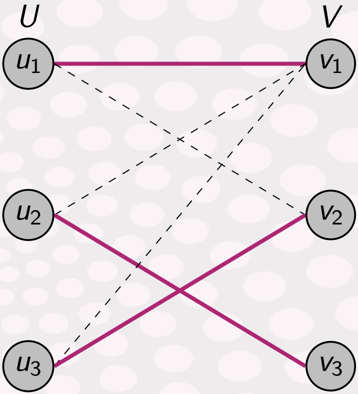


Figure: $OPT = 3$

Competitive ratio

Definition

For $G \in \mathcal{G}$, the competitive ratio is defined as:

$$CR = \frac{\mathbb{E}(ALG(G))}{OPT(G)}$$

Note that $0 \leq CR \leq 1$.

The usual frameworks

- **Adversarial (Adv)**: $\mathcal{G}$ can be any graph. The CR is defined by,

$$CR^{\text{adv}} = \min_{G \in \mathcal{G}} \frac{\mathbb{E}(ALG(G))}{OPT(G)}$$

- **Stochastic (IID)**: The vertices of V are drawn iid from a distribution. (precise definition given latter)

$$CR^{\text{sto}} = \frac{\mathbb{E}(ALG(G))}{OPT(G)}$$

Online matching with unitary budget: Greedy algorithm

Algorithm

For $t = 1, \ldots, |V|$:

    Match v_t to any free neighbor at random

end

Performance of Greedy

Theorem (informal)

In the Adversarial setting, for Greedy (and any deterministic alg.)

$$CR(\text{Greedy}) = \frac{1}{2}$$

A randomized algorithm can achieve,

$$CR(ALG) \geq 1 - \frac{1}{e} \approx 0.63$$

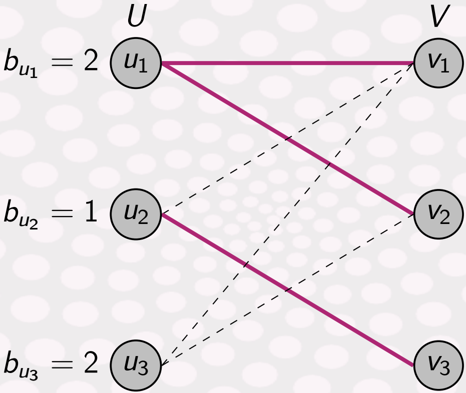
Online b-matching problem: Balance algorithm

The b-matching problem

Problem definition

- Let $G = (U, V, E)$ be a bipartite graph.
- U set of offline nodes, nodes in V are discovered sequentially.
- Each node in U has a budget $b_u > 1$.

The b−matching



The Balance algorithm

Balance algorithm

Algorithm

For $t = 1, \ldots, |V|$:

Match v_t to a neighbor with highest remaining budget.

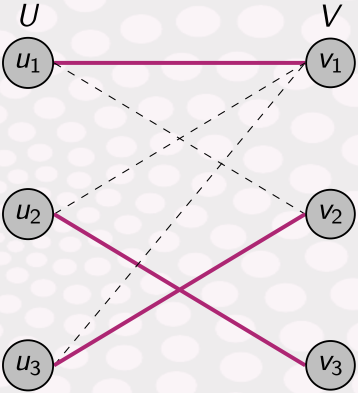
end

Balance algorithm

$b_{u_1,0} = 2$ after match $b_{u_1,1} = 1$

$b_{u_2,0} = 1$ after match $b_{u_2,1} = 0$

$b_{u_3,0} = 2$ after match $b_{u_1,1} = 1$



Performance of Balance

Theorem (informal)

[Kalyanasundaram and Pruhs 2000], when $b_u = b$ for all $u \in U$,

$$CR(\text{Balance}) = 1 - \frac{1}{(1 + 1/b)^b}$$

[Albers and Schubert 2021] with different budget b_u,

$$CR(\text{Balance}) = 1 - \frac{1}{(1 + 1/b_{\min})^{b_{\min}}}, \qquad \text{with } \min_{u \in U} b_u$$

More realistic setting: online matching with budget refills

Problem definition: graph construction

Let $G \in \mathcal{G}$, with $G = (U, V, E)$ a bipartite graph,

- $|U| = n, |V| = T$ with $T \geq n$.
- Nodes in U are offline and nodes in V are revealed sequentially,
- Each node in U has a budget $b_{u,t} \geq 0$ at time $t \in [T]$.

Problem definition: matching construction

A matching on G is a binary matrix $x \in \{0,1\}^{n \times T}$ s.t.

- $\forall (u,t) \in U \times V,\ (u,t) \notin E \Rightarrow x_{u,t} = 0$
  (only edges in E can be matched)
- $\forall t \in V,\ \sum_{u \in U} x_{u,t} \leq 1$
  (no V-node can be matched twice)
- $\forall (u,t) \in U \times V,\ b_{u,t-1} < 1 \Rightarrow x_{u,t} = 0$
  (U-nodes need some positive budget to be matched)

Frameworks considered

We will consider two frameworks for $\mathcal{G}$ and the budget dynamics $b_{u,t}$,

- The **stochastic** framework.
  The graph and the refills are stochastic.
- The **adversarial** framework.
  graph is adversarial, refills deterministic.

The stochastic framework

$\mathcal{G}$ is a family of Erdős–Rényi sparse random graphs:

- Edges occurring independently with probability $p = a/n$.
- Each node in U has a budget $b_{u,t} \in \mathbb{N}$. Budget dynamics:

$$b_{u,t} = \min(K, b_{u,t-1} - x_{u,t} + \eta_t) \quad (1)$$

η_t is a realization of a Bernoulli random variable $\mathcal{B}(\frac{\beta}{n})$.

The matching size created by Greedy

Theorem: (first result)

For $\psi = \frac{T}{n} \geq 1$, with high probability $Greedy(G, T)$ is given by,

$$Greedy(G, T) = nh(\psi) + o(n)$$

where $h(\tau)$ is solution of the following system denoted (A),

$$\begin{cases} \dot{h}(\tau) = 1 - e^{-a(1-z_0(\tau))} & 1/n \leq \tau \leq \psi \\ \dot{z}_0(\tau) = -z_0(\tau)\beta + \frac{z_1(\tau)}{1-z_0(\tau)}(1 - e^{-a+az_0(\tau)}) & \text{for } k = 0 \\ \dot{z}_k(\tau) = (z_{k-1}(\tau) - z_k(\tau))\beta + (z_{k+1}(\tau) - z_k(\tau))\frac{1-e^{-a+az_0(\tau)}}{1-z_0(\tau)} & \text{for } 1 \leq k \leq K-1 \\ \dot{z}_k(\tau) = \beta\, z_{k-1}(\tau) - z_k(\tau)\frac{1-e^{-a(1-z_0(\tau))}}{1-z_0(\tau)} & \text{for } k = K \\ \sum_{k=0}^{K} z_k(\tau) = 1 \end{cases}$$

Moving to the stationary solution

- Solving the ODE satisfied by $h(\tau)$ requires finding $z_0(\tau)$.
- Solving the system of ODE satisfied by z_k is difficult ...
- Focusing on the stability of the stationary solution of the system.

First results for $K = 1$

For $K = 1$, (A) is reduced to ,

$$(A_1) = \begin{cases} \dot{z}_0(t) & = -\beta\, z_0(t) + \frac{z_1(t)}{1 - z_0(t)}\left(1 - e^{-c(1 - z_0(t))}\right) \\ \dot{z}_1(t) & = \beta z_0(t) - \frac{z_1(t)}{1 - z_0(t)}\left(1 - e^{-c(1 - z_0(t))}\right) \\ z_0(t) & + z_1(t) = 1 \end{cases}$$

Stability for $K = 1$

The stationary solution of (A_1) is **exponentially stable** and is given by,

$$z_0^* = \frac{1}{\beta} - \frac{1}{a} W \left(\frac{a}{\beta} e^{-a\left(1-\frac{1}{\beta}\right)} \right)$$

$$z_1^* = z_0^* \frac{\beta}{g(z_0^*)}$$

where W is the Lambert function and $g(z_0^*) = \frac{1-e^{-a(1-z_0^*)}}{1-z_0^*}$.

Result for $K = 1$

Corollary

For $K = 1$ and $\psi = \frac{T}{n} \geq 1$, with high probability ,

$$|Greedy(G, T) - nh^*(\psi)| \leq CT^{1-\epsilon}$$

with $h^*(x) = \int_0^x (1 - e^{-a(1-z_0^*)})\mathrm{d}\tau = x(1 - e^{-a(1-z_0^*)})$.
And $\epsilon > 0$, C is a known constant.

Stability for $K \geq 1$

The stationary solution of (A) is **asymptotically stable** and is given by,

$$z^* = \left(z_0^*, z_0^* \frac{\beta}{g(z_0^*)}, \ldots, z_0^* \left(\frac{\beta}{g(z_0^*)} \right)^K \right)$$

z_0^* is the unique solution of $\sum_{k=0}^{K} z_0^* \left(\frac{\beta}{g(z_0^*)} \right)^k = 1$ with g defined as previously.

Result for $K \geq 1$

Corollary

For $K \geq 1$ and $\psi = \frac{T}{n} \geq 1$, with high probability,

$$|Greedy(G, T) - nh^*(\psi)| \leq o(T)$$

with $h^*(x) = \int_0^x (1 - e^{-a(1-z_0^*)}) \mathrm{d}\tau = x(1 - e^{-a(1-z_0^*)})$, and z_0^* defined as previously.

Convergence of the CR

Theorem(informal)

$$CR(Greedy) \geq \frac{nb_0 + \beta T - \beta T z_0^* \left(\frac{\beta}{g(z_0^*)}\right)^K - n z_0^* \sum_{k=1}^{K} k \left(\frac{\beta}{g(z_0^*)}\right)^k}{nb_0 + \beta T} + O(T^{-1/4})$$

And for $K = 1$, if $\frac{a}{\beta}$ is small enough, then

$$|CR(Greedy) - 1| \leq O(T^{-\frac{1}{4}})$$

For $K \geq 1$, if β is small enough then

$$|CR(Greedy) - 1| \to 0$$

where z_0^* and g defined as previously.

Key takeaways for the stochastic case

- The matching size created by Greedy on the online Erdős- Rényi graph with budget refills is close to the solution of an ODE.
- Based on the stationary solution of the ODE, we have an exact approximation the matching size.
- Under specific conditions on the problem parameters the CR of Greedy converges to 1.

The adversarial framework

$G = (U, V, E)$ is a bipartite graph generated by an oblivious adversary:

- $|U| = n$ and $|V| = T$ with $T \geq n$.
- Each node in U has a budget $b_{u,t} \in \mathbb{N}$. Budget dynamics:

$$b_{u,t} = b_{u,t-1} - x_{u,t} + \mathbb{1}_{t \bmod m = 0} \tag{2}$$

Result when $m \geq \sqrt{T}$

Theorem (informal)

For $m \geq \sqrt{T}$,

$$CR(Balance) \leq 1 - \frac{1}{\left(1 + \frac{1}{b_0}\right)^{b_0}}$$

Result when $m = o(\sqrt{T})$

Theorem (informal)

For $m = o(\sqrt{T})$,

$$CR(balance) \leq \underbrace{1 - \frac{(1-\alpha)}{e^{(1-\alpha)}}}_{\simeq 0.73325\ldots} \tag{3}$$

where α is defined by $\frac{1}{2} = \int_0^{\alpha} \frac{xe^x}{1-x} \mathrm{d}x$.

Balance is the optimal deterministic algorithm

Theorem (informal)

$$\sup_{ALG} \inf_{G \in \mathcal{G}} CR(ALG) \leq \inf_{G \in \mathcal{G}} CR(Balance) \quad (4)$$

.

Summary and future works

Summary

- In the stochastic framework: the matching size of Greedy converges to a function depending on the stationary solution of a system of ODE. And depending on the problem parameters the CR converges to 1.
- In the adversarial framework: depending on the refill frequency we get upper bounds on the CR of Balance algorithm.

Summary and future works

Future works

- Prove that for $K > 1$ there is exponential stability of z^*.
- Lower bound of the CR of Balance.

Thank you.

Bibliography I

- Albers, Susanne and Sebastian Schubert (2021). "Optimal Algorithms for Online b-Matching with Variable Vertex Capacities". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021)*. Ed. by Mary Wootters and Laura Sanità. Vol. 207. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2:1–2:18.
- Kalyanasundaram, Bala and Kirk R. Pruhs (2000). "An optimal deterministic algorithm for online b-matching". In: *Theoretical Computer Science* 233.1, pp. 319–325.
- Wormald, Nicolas (1999). *The differential equation method for random graph processes and greedy algorithms.* URL: https://users.monash.edu.au/~nwormald/papers/de.pdf.

Sketch of Proof

- Let $Y_k(t)$ be the number of nodes in U with budget $k \geq 0$.
- $Greedy(G, t)$ the matching size obtained by GREEDY on the online Erdös-Rényi bipartite graph with budget refills at time t.
- Since the evolution of $Greedy(G, t)$ depends on $Y_0(t)$, the idea is to prove that $(Y_k(t))_{0 \leq k \leq K}$ is close to the solution of a system of an ODE using the differential equation method.
- Then, do the same for $Greedy(G, t)$.

The matching size at time $t+1$ is defined as follows,

$$M(t+1) = M(t) + \mathbb{1}_{\{x_{u,t+1}=1,\ u \in U_k(t+1)\}}$$

Moving to conditional expectation we get,

$$\mathbb{E}\left[M(t+1) - M(t) | M(t)\right] = \mathbb{P}\left(x_{u,t+1} = 1,\ u \in U_k | M(t)\right)$$
$$= 1 - \left(1 - \frac{a}{N}\right)^{N - Y_0(t)}$$

M(t) **depends on** $Y_0(t)$!

Applying the differential equation method on $(Y_k(t))_{k\geq 0}$

Using Wormald 1999 results, we have $Y_k(t) = n\, z_k(t/n) + O(\lambda\, n)$ with probability $1 - \mathcal{O}(\frac{\gamma}{\lambda} \exp - \frac{n\lambda^2}{\gamma^3})$ with $\gamma = 3n$, $\lambda = an^{-1/4}$, where z_k is solution of the following system, $\forall \tau \in [0, 1]$,

$$\dot{z}_0(\tau) = -z_0(\tau)\beta + \frac{z_1(\tau)}{1 - z_0(\tau)}(1 - e^{-a+az_0(\tau)}) \qquad \text{for } k = 0$$

$$\dot{z}_k(\tau) = (z_{k-1}(\tau) - z_k(\tau))\beta + (z_{k+1}(\tau) - z_k(\tau))\frac{1 - e^{-a+az_0(\tau)}}{1 - z_0(\tau)} \quad \text{for } k \geq 1$$

$$\sum_{k\geq 0} z_k(\tau) = 1$$

Applying differential equation method to $M(t)$

Using Wormald 1999 results, we have $M(t) = n\,h(t/n) + O(\lambda_m\, n)$ with probability $1 - \mathcal{O}(\frac{\gamma_m}{\lambda_m} \exp -\frac{n\lambda_m^2}{\gamma_m^3})$, where $\gamma_m = 1$, $\lambda_m = an^{-1/4}$, where h is solution of the following equation,

$$\dot{h}(\tau) = 1 - e^{-a(1-z_0(\tau))}$$