A Two-Player Singleton Stochastic Congestion Game with Asymmetric Information

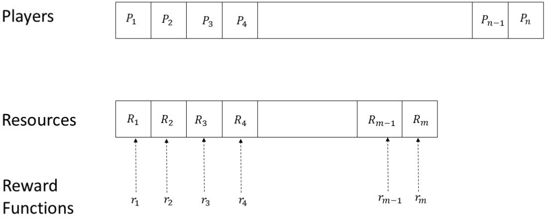
March 17, 2025

Table of Contents

Outline

Congestion Games

- A Congestion Game[1] is a tuple $(\mathcal{N}, \mathcal{R}, \mathcal{A}, \boldsymbol{r})$, where,
  - $\mathcal{N} = \{P_1, P_2, ..., P_n\}$ is a set of n players
  - $\mathcal{R} = \{R_1, R_2, ..., R_m\}$ is a set of m resources
  - $\boldsymbol{r} = (r_1, ..., r_m)$, $r_k : \mathbb{N}_0 \mapsto \mathbb{R}$ is the reward function of resource k.
  - Each player chooses a subset of the resources $a_i \in \mathcal{A}_i \subset 2^{\mathcal{R}} \backslash \{\varnothing\}$.
  - $\mathcal{A} = \mathcal{A}_1 \times .... \times \mathcal{A}_n$

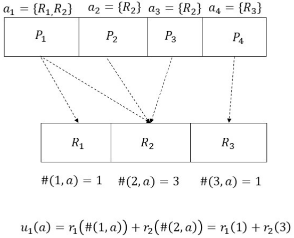


[1] Robert W. Rosenthal. "A class of games possessing pure-strategy Nash equilibria". In: *International Journal of Game Theory* (1973).

Congestion Games Continued

- Action profile $a = (a_1, a_2, ..., a_n)$
- Count function: $\# : R \times A \mapsto \mathbb{N}$, $\#(k, a)$ = number of players choosing k under the action profile a.
- Reward functions r_k are typically non-decreasing
- $r_k(\#(k, a))$ is the per player reward of a resource.
- The utility of player i is,

$$u_i(a) = \sum_{k \in a_i} r_k(\#(k, a)) \tag{1}$$



Congestion Games Properties

- Congestion games fall under potential games[2]
- There are many variants of the game,
  - Weighted Congestion Games[3]
  - Singleton Congestion Games[4]
  - Resource failure[5]
  - Time varying Dynamic/Stochastic settings[6]

[2] Dov Monderer and Lloyd S. Shapley. "Potential Games". In: *Games and Economic Behavior* (1996).

[3] Kshipra Bhawalkar, Martin Gairing, and Tim Roughgarden. "Weighted Congestion Games: Price of Anarchy, Universal Worst-Case Examples, and Tightness". In: *Lecture Notes in Computer Science*. 2010.

[4] Dimitris Fotakis et al. "The structure and complexity of Nash equilibria for a selfish routing game". In: *Theoretical Computer Science* (2009).

[5] Jinhuan Wang, Kaichen Jiang, and Yuhu Wu. "On congestion games with player-specific costs and resource failures". In: *Automatica* (2022).

[6] Martin Hoefer et al. "Competitive routing over time". In: *Theoretical Computer Science* (2011); Haris Angelidakis, Dimitris Fotakis, and Thanasis Lianeas. "Stochastic Congestion Games with Risk-Averse Players". In: *Lecture Notes in Computer Science*. 2013.

Applications of Congestion Games

- The Applications include,
  - Service chain composition[7]
  - Network design[8]
  - Load balancing[9]
  - Spectrum sharing[10]
  - Radio access selection[11]
  - Modelling the Migration of species[12]

---

[7] Shuting Le, Yuhu Wu, and Mitsuru Toyoda. “A Congestion Game Framework for Service Chain Composition in NFV with Function Benefit”. In: *Inf. Sci.* (2020).

[8] E. Anshelevich et al. “The price of stability for network design with fair cost allocation”. In: *45th Annual IEEE Symposium on Foundations of Computer Science.* 2004.

[9] Ioannis Caragiannis et al. “Tight Bounds for Selfish and Greedy Load Balancing”. In: *Algorithmica* (2006).

[10] Sahand Ahmad et al. *Spectrum Sharing as Network Congestion Games*. 2009.

[11] Marc Ibrahim, Kinda Khawam, and Samir Tohme. “Congestion Games for Distributed Radio Access Selection in Broadband Networks”. In: *2010 IEEE Global Telecommunications Conference GLOBECOM 2010.* 2010.

[12] Thomas Quint and Martin Shubik. “A model of migration”. In: (1994).

Our Contributions

What happens when the players,

- have asymmetric information?
- do not trust each other?

Model

- Two players A and B and n resources.
- Each player can choose exactly one resource (Singleton).
- Each resource i has a Reward random variable W_i.
  - If both players choose the same resource i the utility of each player is $W_i/2$.
  - If players choose different resources they get the full reward of the resource.
- W_i are assumed to be independent.
- Reward of player A is,

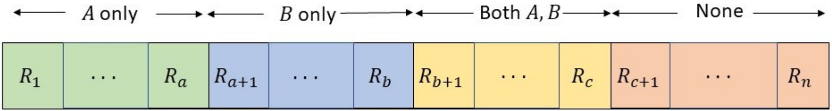
$$R_A = \sum_{k=1}^{n} W_k \mathbb{1}_{(\alpha^A = k)} \left(1 - \frac{\mathbb{1}_{(\alpha^B = k)}}{2}\right)$$

$$\alpha^X = \text{ resource choosen by X} \tag{2}$$

Outline

Information Asymmetry

- Both players know the distribution $\boldsymbol{W} = (W_1, W_2, ..., W_n)$
- Each player observes the realization of the reward random variable of some of the resources.
- $\boldsymbol{W} = (\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{V})$
  - $\boldsymbol{X} \in \mathbb{R}^a$ - Only A
  - $\boldsymbol{Y} \in \mathbb{R}^{b-a}$ - Only B
  - $\boldsymbol{Z} \in \mathbb{R}^{c-b}$ - Both A and B
  - $\boldsymbol{V} \in \mathbb{R}^{n-c}$ - None
- $0 \leqslant a \leqslant b \leqslant c \leqslant n$
- We will say **A sees the resource** i for $1 \leqslant i \leqslant a$ or $b + 1 \leqslant i \leqslant c$.



Expected Reward

- In our analysis we fix $\boldsymbol{Z}$. Let $E_k = \mathbb{E}\{W_k|\boldsymbol{Z}\}$
- The expected reward of player A can be simplified as,

$$\mathbb{E}\{R_A|\boldsymbol{Z}\} = \underbrace{\sum_{k=1}^{a} q_k^A + \sum_{k=a+1}^{n} E_k p_k^A}_{\hat{A} \text{ (Depends only on A's strategy)}} - \underbrace{\frac{1}{2}\left(\sum_{k=1}^{a} q_k^A p_k^B + \sum_{k=a+1}^{b} p_k^A q_k^B + \sum_{k=b+1}^{n} E_k p_k^A p_k^B\right)}_{\hat{C}} \quad (3)$$

- For $1 \leqslant k \leqslant n$,

$$p_k^A = \mathbb{E}\{\mathbb{1}_{(\alpha^A=k)}|\boldsymbol{Z}\}, \quad p_k^B = \mathbb{E}\{\mathbb{1}_{(\alpha^B=k)}|\boldsymbol{Z}\}$$
$$q_k^A = \mathbb{E}\{W_k\mathbb{1}_{(\alpha^A=k)}|\boldsymbol{Z}\}, \quad q_k^B = \mathbb{E}\{W_k\mathbb{1}_{(\alpha^B=k)}|\boldsymbol{Z}\} \quad (4)$$

Expected Reward Continued

- Similarly for player B, we have,

$$\mathbb{E}\{R_B|\boldsymbol{Z}\} = \underbrace{\sum_{k=1}^{a} E_k p_k^B + \sum_{k=a+1}^{b} q_k^B + \sum_{k=b+1}^{n} E_k p_k^B}_{\hat{B}} \\ \underbrace{- \frac{1}{2}\left(\sum_{k=1}^{a} q_k^A p_k^B + \sum_{k=a+1}^{b} p_k^A q_k^B + \sum_{k=b+1}^{n} E_k p_k^A p_k^B\right)}_{\hat{C}} \quad (5)$$

Finding a Nash Equilibrium

- Finding the best response of a player for a fixed strategy of the opponent
- Finding a global potential function
- Running the iterated best response algorithm.

Best Response

Recall

- Recall for $1 \leqslant k \leqslant n$,

$$p_k^B = \mathbb{E}\{\mathbb{1}_{(\alpha^B = k)} | \boldsymbol{Z}\} \quad q_k^B = \mathbb{E}\{W_k \mathbb{1}_{(\alpha^B = k)} | \boldsymbol{Z}\} \tag{6}$$

- For a fixed strategy of B, p_k^B and q_k^B will be fixed.

- The best response of A for a fixed strategy of B is given by $\alpha^A = \arg\max_{1 \leqslant k \leqslant n} A_k$, where A_k is given by,

$$A_k = \begin{cases} W_k \left(1 - \frac{1}{2} p_k^B\right) & \text{if } 1 \leqslant k \leqslant a \\ E_k - \frac{1}{2} q_k^B & \text{if } a + 1 \leqslant k \leqslant b \\ E_k \left(1 - \frac{1}{2} p_k^B\right) & \text{if } b + 1 \leqslant k \leqslant n \end{cases}. \tag{7}$$

- The best response of B can be calculated similarly

Potential Functions

- A potential function f has the following properties,
  - f Depends on the policies of A and B
  - When A changes strategy while B stays in the strategy the change of $\mathbb{E}\{R_A|\boldsymbol{Z}\}$ is equal to the change of f
  - Same is true when B changes strategy while A stays
- **Potential function has to be global. Not player specific!!!**

Potential Function for our Game

Recall

$$\mathbb{E}\{R_B|\boldsymbol{Z}\} = \hat{B} - \hat{C}$$
$$\mathbb{E}\{R_A|\boldsymbol{Z}\} = \hat{A} - \hat{C} \quad (8)$$

- $\hat{B}$ - depends only on B's strategy
- $\hat{A}$ - depends only on A's strategy

- It turns out that our game has an exact potential function,
- The potential function

$$H(A,B) = \hat{A} + \hat{B} - \hat{C}$$
$$= \mathbb{E}\{R_A|\boldsymbol{Z}\} + \underbrace{\hat{B}}_{\text{Does not change when A individually changes strategy}}$$

Iterated Best Response

- Players A and B can iteratively find the best response
- In each iteration
  - First A finds the best response while B stays in the strategy - Potential function $\uparrow$
  - Then B finds the best response while A stays in the strategy - Potential function $\uparrow$
- Potential function will be non-decreasing in each step
- Potential function is bounded.
- Convergence to ϵ-pure Nash equilibrium in at most,

$$\frac{\sum_{k=1}^{n} E_k}{\epsilon}, \tag{9}$$

iterations

Outline

Motivation

- What happens when players,
  - Do not have information regarding opponents' strategy?
  - Do not trust each other?
- Solution: Maximizing the worst-case expected reward.

Steps involved

- We will focus on finding a worst-case strategy for A.
- The steps involved,
  - Fix A's strategy
  - Find the strategy of B which minimizes $\mathbb{E}\{R_A|\boldsymbol{Z}\}$ for the A's fixed strategy
  - Maximize $\mathbb{E}\{R_A|\boldsymbol{Z}\}$ in this case.

Worst-Case Expected Reward

- Recall, when we fix A's strategy p_k^A's and q_k^A's will be fixed.
- The strategy of B which minimizes $\mathbb{E}\{R_A|\boldsymbol{Z}\}$ is given by,

$$\alpha^B = \arg \max_{1 \leqslant k \leqslant n} \mu_k,$$

where

$$\mu_k = \begin{cases} q_k^A & \text{if } 1 \leqslant k \leqslant a \\ W_k p_k^A & \text{if } a+1 \leqslant k \leqslant b \\ E_k p_k^A & \text{if } b+1 \leqslant k \leqslant n \end{cases} \tag{10}$$

- The worst case expected reward of A is given by,

$$R_{\text{worst}} = \sum_{k=1}^{a} q_k^A + \sum_{k=a+1}^{n} E_k p_k^A - \frac{1}{2} \sum_{k=1}^{n} \mathbb{E}\{\max\{\mu_k\}_{k=1}^{n} |\boldsymbol{Z}\} \tag{11}$$

Maximizing the Worst-Case Expected Reward

- Problem

$$
\begin{aligned}
(P1:)\ \text{maximize} \quad & \sum_{k=1}^{a} q_k^A + \sum_{k=a+1}^{n} E_k p_k^A - \frac{1}{2}\sum_{k=1}^{n} \mathbb{E}\{\max\{\mu_k\}_{k=1}^{n} | \boldsymbol{Z}\} \\
\text{subject to} \quad & p_k^A = \mathbb{E}\{\mathbb{1}_{(\alpha^A = k)} | \boldsymbol{Z}\} \text{ for } 1 \leqslant k \leqslant n, \\
& q_k^A = \mathbb{E}\{W_k \mathbb{1}_{(\alpha^A = k)} | \boldsymbol{Z}\} \text{ for } \text{ for } 1 \leqslant k \leqslant n, \\
& \mu_k = \begin{cases} q_k^A & \text{if } 1 \leqslant k \leqslant a \\ W_k p_k^A & \text{if } a+1 \leqslant k \leqslant b \\ E_k p_k^A & \text{if } b+1 \leqslant k \leqslant n \end{cases}
\end{aligned}
$$

- Two approaches,
  - Direct-approach
  - Drift-plus penalty-based approach

Direct Approach

Recall

$$
\begin{aligned}
(P1 :)\ \text{maximize} \quad & \sum_{k=1}^{a} q_k^A + \sum_{k=a+1}^{n} E_k p_k^A - \frac{1}{2}\sum_{k=1}^{n} \mathbb{E}\{\max\{\mu_k\}_{k=1}^{n} | \boldsymbol{Z}\} \\
\text{subject to} \quad & p_k^A = \mathbb{E}\{\mathbb{1}_{(\alpha^A = k)} | \boldsymbol{Z}\} \text{ for } 1 \leqslant k \leqslant n, \\
& q_k^A = \mathbb{E}\{W_k \mathbb{1}_{(\alpha^A = k)} | \boldsymbol{Z}\} \text{ for } \text{ for } 1 \leqslant k \leqslant n, \\
& \mu_k = \begin{cases} q_k^A & \text{if } 1 \leqslant k \leqslant a \\ W_k p_k^A & \text{if } a+1 \leqslant k \leqslant b \\ E_k p_k^A & \text{if } b+1 \leqslant k \leqslant n \end{cases}
\end{aligned}
$$

1. Finding the region $\mathcal{G} \subset \mathbb{R}^{2n}$ achieved by $(q_1^A, ..., q_n^A, p_1^A, ..p_n^A)$
2. Solving (P1) as a problem in $\mathbb{R}^{2n}$, for $(q_1, ..., q_n, p_1, ..p_n) \in \mathcal{G}$
3. Finding a strategy satisfying the found optimal $(q_1^*, ..., q_n^*, p_1^*, ..p_n^*)$.

$$
p_k^* = \mathbb{E}\{\mathbb{1}_{(\alpha^* = k)} | \boldsymbol{Z}\} \quad q_k^* = \mathbb{E}\{W_k \mathbb{1}_{(\alpha^* = k)} | \boldsymbol{Z}\} \tag{12}
$$

Direct Approach Continued

Recall the steps:

1. Finding the region $\mathcal{G} \subset \mathbb{R}^{2n}$ achieved by $(q_1^A, ..., q_n^A, p_1^A, ..p_n^A)$
2. Solving (P1) as a problem in $\mathbb{R}^{2n}$, for $(q_1, ..., q_n, p_1, ..p_n) \in \mathcal{G}$
3. Finding a strategy satisfying the found optimal $(q_1^*, ..., q_n^*, p_1^*, ..p_n^*)$.

$$p_k^* = \mathbb{E}\{\mathbb{1}_{(\alpha^*=k)}|\boldsymbol{Z}\} \quad q_k^* = \mathbb{E}\{W_k \mathbb{1}_{(\alpha^*=k)}|\boldsymbol{Z}\} \tag{13}$$

- The first step can be done easily when $a = 1$ [13].
- For $a = 1$, the second step can be done using the Stochastic sub-gradient method.
- Third step finds a mixture of **threshold strategies** for $a = 1$.

[13] When $a = 1$,

$$\mathcal{G} = \{(q, \boldsymbol{p}) | p_k \in [0, 1], \sum_{k=1}^{n} p_k = 1, q \in \mathbb{R}, \mathbb{E}\{W_1 | F_{W_1}(W_1) \leqslant p_1^A\} p_1 \leqslant q \leqslant \mathbb{E}\{W_1 | F_{W_1}(W_1) \geqslant 1 - p_1\} p_1\} \tag{14}$$

Threshold Strategy

- A threshold strategy is defined by $\boldsymbol{C} \in \mathbb{R}^n$ and is given by,

$$\alpha^A = \arg \max_{1 \leqslant k \leqslant n} \{\{C_j W_j\}_{j=1}^{a}, \{C_j\}_{j=a+1}^{n}\}, \tag{15}$$

- It turns out that the direct method finds a mixture of threshold strategies when $a = 1$.
- Can we find such a mixture for the general case?

Drift-plus Penalty Approach

- Idea is to find T threshold strategies $\boldsymbol{C}(t)$ for $1 \leqslant t \leqslant T$, which can be used in a mixture.
- Recall that

$$\alpha^A(t) = \arg \max_{1 \leqslant k \leqslant n} \{\{C_j(t) W_j\}_{j=1}^{a}, \{C_j(t)\}_{j=a+1}^{n}\} \tag{16}$$

- Can be done using treating $C_j(t)$'s as virtual queues.
- Can be used for the general case
- Slower compared to the direct method

Algorithm

- Choose a parameter V and initialize $\boldsymbol{C}(0) = 0$
- For each iteration $t \in \{0, 1, 2, ..., T-1\}$
  1. Sample $\boldsymbol{X}(t)$
  2. **(P2)**: Choose $\gamma(t)$ to solve,

$$(P2) : \underset{\gamma(t)}{\text{maximize}} \quad Vf(\gamma(t)) - \sum_{j=1}^{n} C_j(t)\gamma_j(t) \tag{17a}$$

$$\text{subject to} \quad \gamma(t) \in \left(\prod_{j=1}^{a}[0, E_j]\right) \times [0,1]^{n-a} \tag{17b}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is given by,

$$f(\boldsymbol{x}) = \sum_{j=1}^{a} x_j + \sum_{j=a+1}^{n} E_j x_j - \frac{1}{2}\mathbb{E}\{\max\{\boldsymbol{x}_{1:a}, \{x_j W_j\}_{j=a+1}^{b}, \{x_j E_j\}_{j=b+1}^{n}\}|\boldsymbol{Z}\}$$

  3. Choose action $\alpha^A(t)$ using the threshold strategy $\boldsymbol{C}(t)$.
  4. Updates the virtual queues using,

$$\begin{aligned} C_j(t+1) &= \max\{C_j(t) + \gamma_j(t) - X_j(t)\mathbb{1}_{\alpha^A(t)=j}, 0\}, \forall 1 \leqslant j \leqslant a, \\ C_j(t+1) &= \max\{C_j(t) + \gamma_j(t) - \mathbb{1}_{\alpha^A(t)=j}, 0\}, \forall a+1 \leqslant j \leqslant n. \end{aligned} \tag{18}$$

Solving (P2)

Recall

$$(P2): \underset{\gamma(t)}{\text{maximize}} \quad Vf(\gamma(t)) - \sum_{j=1}^{n} C_j(t)\gamma_j(t) \tag{19a}$$

$$\text{subject to} \quad \gamma(t) \in \left(\prod_{j=1}^{a}[0, E_j]\right) \times [0,1]^{n-a} \tag{19b}$$

- Can be solved using the stochastic subgradient method with projections onto the feasible set[14]
- Can be sped up in certain cases for instance when $C_j(t) \notin [Vv_j/2, Vv_j]$ where,

$$v_j = \begin{cases} 1 & \text{if } 1 \leqslant k \leqslant a \\ E_k & \text{otherwise} \end{cases} \tag{20}$$

[14] Stephen Boyd and Almir Mutapcic. "Stochastic subgradient methods". In: *Lecture*

Performance Evaluations

- The Expected reward of the algorithm generated by the algorithm is bounded as,

$$\mathbb{E}\{R^{\text{mixed}}|\boldsymbol{Z}\} \geqslant f^{\text{opt}} - \frac{D}{V} - \frac{3v_{\max}}{2}\sqrt{\frac{2n(D + V(f^{\max} - f^{\text{opt}}))}{T}}, \quad (21)$$

where

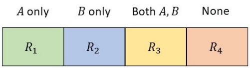
- $D = n - a + \frac{1}{2}\sum_{j=1}^{a}(E_j^2 + \mathbb{E}\{W_k(t)^2\})$
- f^{opt} is the optimal value of (P1)
- $f^{\max} = \sup_{x \in \left(\prod_{j=1}^{a}[0,E_j]\right) \times [0,1]^{n-a}} f(x)$
- $v_{\max} = \max\{\{v_j\}_{j=1}^{n}\}$, where

$$v_k = \begin{cases} 1 & \text{if } 1 \leqslant k \leqslant a \\ E_k & \text{otherwise} \end{cases} \quad (22)$$

Outline

E_1 vs $\mathbb{E}\{R_A|Z\}$

- W_j are exponential, $E_2 = \mathbb{E}\{W_3\} = E_4 = 1$ and $a, b, c, n = 1, 2, 3, 4$
- For the first three cases B is playing the greedy strategy



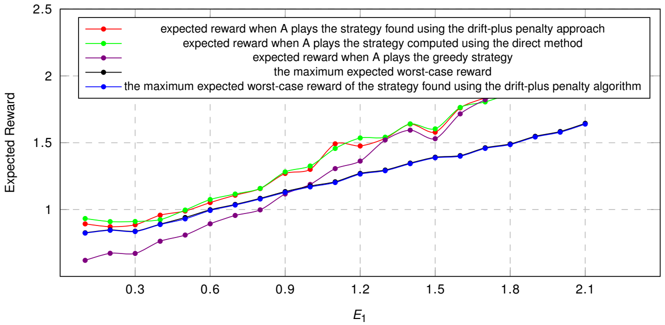


Figure: The expected reward of A vs E_1

Outline

Conclusions

We have considered the Two-Player Singleton Stochastic Congestion Game with Asymmetric Information.

- Existence of an exact potential function
- An iterative best response algorithm to find the ϵ-pure Nash equilibrium
- Two algorithms to maximize the worst-case expected reward of the first player.
  - Direct method: has lower computational complexity but can only be applied in specific cases.
  - Drift-plus penalty algorithm: can be applied in the general scenario but has high computational complexity.

Thank You