University of Twente Summer Internship Project 2019

FINAL REPORT

Upper Bound of Price of Anarchy in Content Placement Game

Author:
Jiduan Wu

Supervisors:
Dr.A.Skopalik(UT)
Dr.ir.J.Goseling(UT)

2019 July 1st - August 27th

Contents

1 Introduction

Our story happens in the world of cellular networks, there are users who search files online like we normally do; base stations who cache the files with finite capacity to provide users within its covering region with online service. As the ascending of data traffic is expected to explode in the near future. To utilize the limited resource: backhaul links that connect base stations to the core network as efficiently as possible, a natural thought for each base station is caching the "proper" files with precious capacity.

From a game theoretical perspective, we treat base stations as selfish players, they choose which files they cache and we call the different strategy combinations as strategy profiles. We simplify the real situation and use the model with only one user who searches for one file each time. In this report we use Zipf distribution as the probability of the user requests different files. We use miss probability: *the probability that a user can't get the file she/he wants* to evaluate the performance of each strategy profile, all of these elements make up the *content placement game* model, which is introduced in [1]

The main goal of this project is to obtain an upper bound on price of anarchy in content placement game, ideal results on price of anarchy have been observed through simulations in [1], which gives us the hope that we can find a "good" upper bound on the price of anarchy theoretically. The following results and analysis are all based on this objective.

The contributions in this report are given as follows:

- Upper bound of PoA under some assumptions in content placement game, 2-player case and grid case.
- Four lemmas to describe the relationship between Nash equilibrium and optimal strategy profile.

The organization of this report is given as follows:

- Firstly, we give a detailed description on the content placement game and notations, terms we use.
- Then we explain the main method we adopt to solve the problem.
- We present the lemmas and results and corresponding proof.
- Lastly, we summarize the work unfinished and make a brief summary of this project.

2 Model Description: Content Placement Game

Here are the key elements in our game:

- N base stations are the players in our game.
  - All the base stations have the same capacity, they can cache at most k files.
  - Each user s and the set of all the users Φ
  - Each of them covers a disc region with constant area A.
  - These regions can intersect with each other arbitrarily.
  - We consider regions which are covered by different sets of players differently, you can have a better grasp of this through the nice picture figure 1.

$$A_s = (\cap_l \in s\bar{A}_l) \cap (\cap_{l \notin s} \bar{A}_l^c), p_s = \frac{|A_s|}{\cup_{s\in\Phi} A_s}$$

- There are J files in total, denoted as $\{1, 2, ..., J\}$, from now on we assume $J > N \times k$.
  - We use Zipf distribution to describe the probability a_j of the user requesting for file j

$$a_j = \frac{j^{-\gamma}}{\sum_{j=1}^{J} j^{-\gamma}}$$

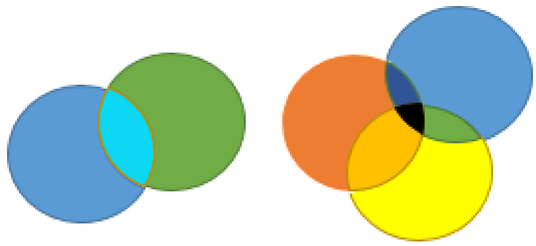


Figure 1: Different colors denote different regions

The choices of players on which k files to cache form a strategy profile, we use an indicator vector $b^{(i)}$ or s_i to denote the strategy of each player i and matrix $B = (b^{(1)}, ..., b^{(N)})$ or s to denote a strategy profile.

$$\begin{cases} b_j^{(i)} = 1, \text{when player i caches file j} \\ b_j^{(i)} = 0, \text{when player i doesn't cache file j} \end{cases}$$

- Cost function $C_i(\mathbf{s})$ for each player is the probability that the user might not be satisfied within the region covered by the player.

$$c_i(s) = \sum_{j=1}^{J} a_j (1 - b_j^{(i)}) \sum_{A \in \Theta, i \in A} p_A \prod_{l \in A \setminus \{i\}} (1 - b_j^{(l)}) \tag{1}$$

where Θ is the power set of [1:N], A is a subset of base stations, s is an arbitrary strategy profile.

- Social Cost Function: the probability that the user does not find the requested file.

$$c(s) = f(B) = \sum_{j=1}^{J} a_j \sum_{s \in \Phi} p_s \prod_{l \in s} (1 - b_j^{(l)}) \tag{2}$$

3 Main Theoretical Methods Applied

Price of Anarchy, which describes the worst outcome of selfish strategy compared to the optimal solution, has been studied extensively as an essential conception in game theory, one of the most famous method is *smoothness* construed by Tim Roughgarden [2].

The main idea of theorems is based on a derivative inequality of classical smoothness inequality.

$$\lambda c(s^*) \geq \sum_{i=1}^{N} (c_i(s_i^*, s_{-i}) - c_i(s)) + (1 - \mu) c(s) \tag{3}$$

where s is the Nash equilibrium and s^* is the optimal solution, by bounding $\sum_{i=1}^{N} c_i(s_i^*, s_{-i})$ with terms of $c(s^*), c(s)$ and apply the definition of Nash equilibrium: no players can benefit from unilateral deviation, we obtain the classical smoothness inequality below

$$\lambda c(s^*) \geq \sum_{i=1}^{N} (c_i(s_i^*, s_{-i}) - c_i(s)) + (1 - \mu) c(s) \geq (1 - \mu) c(s) \tag{4}$$

Hopefully we will get an upper bound $\frac{\lambda}{1-\mu}$ of PoA, here λ, μ may depend on the capacity, the intersecting area and so on.

Four lemmas are mainly obtained through observing simple cases, two or three base stations with small capacity such as two or three, guessing and then proving.

4 Results

4.1 Preliminary result

Definition 4.1. We introduce the conception of miss probability density ρ_T in region T

$$\rho_T = \frac{\sum_i a_i}{\text{The whole area}}$$

where i is summed over the set of files that can not be accessed in region T, and the whole area equals $2 - A$.

Remark. The reason we introduce the conception of miss probability density is to keep the proof short and clear, and the word "density" can properly describe the fact that the user have accesses to more files within the intersecting area thus is less likely to miss the file requested.

Lemma 1. *In 2-player case, assume the intersecting area is A, the following inequality holds.*

$$c(s) \leq \sum_{i=1}^{2} c_i(s) \leq \frac{2}{2-A} c(s) \tag{5}$$

Remark. From the later part you will see it's the lower bound that matters.

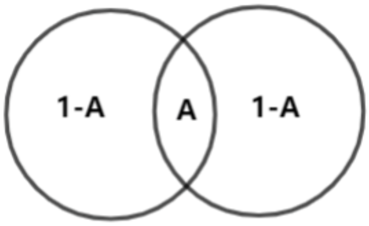


Figure 2: 2-player case

Proof. We denote the regions from left to right as region 1,2,3 and the left player as player 1, the right player as player 2. Let ρ_1, ρ_2, ρ_3 denote the density of miss probability in region 1, 2, 3, we can easily find that $\rho_2 \leq min\{\rho_1, \rho_3\}$, this proof only involves simple computations.

$$c(s) = (\rho_1 + \rho_3)(1 - A) + \rho_2 A \tag{6}$$

$$c_1(s) = \rho_1(1 - A) + \rho_2 A \tag{7}$$

$$c_2(s) = \rho_3(1 - A) + \rho_2 A \tag{8}$$

$$\sum_{i=1}^{2} c_i(s) - c(s) = \rho_2 A \geq 0 \tag{9}$$

$$\frac{2}{2-A} c(s) - \sum_{i=1}^{2} c_i(s) = \frac{A(1-A)(\rho_1 + \rho_3 - 2\rho_2)}{2-A} \geq 0 \tag{10}$$

□

Using the definition of miss probability density and *lemma 1*, we present the first result we obtained as *theorem 4.1.*

Theorem 4.1. *In two-player case, assume the intersecting area is A, if $A < \frac{1}{2}$, then $\frac{1-A}{1-2A}$ is an upper bound on PoA.*

Proof. Our objective is to construe the derivative inequality

$$\lambda c(s^*) \geq \sum_{i=1}^{2}(c_i(s_i^*, s_{-i}) - c_i(s)) + (1-\mu)c(s)$$

where s^* is the optimal strategy and s is an arbitrary Nash equilibrium. Using the miss probability density conception, we can clearly express the cost function for each player and social cost function.

$$\begin{aligned} c_1(s_1^*, s_2) + c_2(s_1, s_2^*) &= \rho_1^*(1-A) + (\rho_1^* \cap \rho_3)A + \rho_3^*(1-A) + (\rho_1 \cap \rho_3^*)A \\ &= (\rho_1^* + \rho_3^*)(1-A) + (\rho_1^* \cap \rho_3 + \rho_1 \cap \rho_3^*)A \end{aligned} \tag{11}$$

where $\rho_1^* \cap \rho_3, \rho_1 \cap \rho_3^*$... denote the "mixed" miss probability density after deviations of players, and we should note that $\rho_i \cap \rho_j \leq min\{\rho_i, \rho_j\}$ for any i, j.

$$c(s) = (\rho_1 + \rho_3)(1-A) + \rho_2 A \tag{12}$$

$$c(s^*) = (\rho_1^* + \rho_3^*)(1-A) + \rho_2^* A \tag{13}$$

And note the common term $(\rho_1^* + \rho_3^*)(1-A)$ (this term corresponds with regions covered only by one base station), we can easily get

$$c_1(s_1^*, s_2) + c_2(s_1, s_2^*) \leq \frac{A}{1-A}c(s) + c(s^*) \tag{14}$$

Apply lemma 1, we have

$$\sum_{i=1}^{2}(c_i(s_i^*, s_{-i}) - c_i(s)) \leq c(s^*) + (\frac{A}{1-A} - 1)c(s) \tag{15}$$

$$\sum_{i=1}^{2}(c_i(s_i^*, s_{-i}) - c_i(s)) + (1 - \frac{A}{1-A})c(s) \leq c(s^*) \tag{16}$$

This inequality has the same form as the derivative inequality, let $\lambda = 1, 1-\mu = 1 - \frac{A}{1-A}$, we get

$$\frac{\lambda}{1-\mu} = \frac{1}{1-\frac{A}{1-A}} = \frac{1-A}{1-2A} \tag{17}$$

And we can see from the formula that when $A \to 0$, $\frac{\lambda}{1-\mu} \to 1$, which corresponds with our intuition. ☐

4.2 Find out more

4.2.1 More Players: Grid Model

Assume base stations are fixed on a grid with $m \times n$ vertices, they intersect with their neighbors pairwise. We still use A to denote the intersecting area (intersecting regions have the same area in this case).

We use the coordination $(i, j), 1 \leq i \leq m, 1 \leq j \leq n$ to denote each player, and we try to use "circles" to deal with boundaries, which means when $i = 1$, $i - 1 = m + 1$, when $j = 1, j - 1 = n + 1$, when $i = m, i + 1 = 1$, when $j = n, j + 1 = 1$.

Theorem 4.2. *For a grid model described above with $m \times n$ players, $\frac{4A}{1-8A}$ is an upper bound on PoA.*

Proof. There is no essential difference in cost functions between the grid case and 2-player case.

$$c(s) = \sum_{i,j} \rho_{i,j}(1-4A) + \sum_{i,j}(\rho_{i-1,j} \cap \rho_{i,j})A + \sum_{i,j}(\rho_{i,j-1} \cap \rho_{i,j})A \tag{18}$$

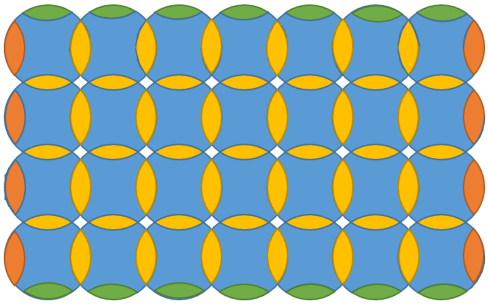


Figure 3: Grid Model

$$c(s^*) = \sum_{i,j} \rho^*_{i,j}(1-4A) + \sum_{i,j}(\rho^*_{i-1,j} \cap \rho^*_{i,j})A + \sum_{i,j}(\rho^*_{i,j-1} \cap \rho^*_{i,j})A \tag{19}$$

$$c_{i,j}(s^*_{i,j}, s_{-(i,j)}) = \rho^*_{i,j}(1-4A) + (\rho_{i-1,j} \cap \rho^*_{i,j} + \rho_{i+1,j} \cap \rho^*_{i,j} + \rho_{i,j+1} \cap \rho^*_{i,j} + \rho_{i,j-1} \cap \rho^*_{i,j})A \tag{20}$$

Note $\rho_1 \cap \rho_2 \leq min\{\rho_1, \rho_2\}$ still holds and each player has four neighbors, we can find the inequality below hold.

$$\sum_{i,j} c_{i,j}(s^*_{i,j}, s_{-(i,j)}) \leq c(s^*) + \frac{4A}{1-4A}c(s) \tag{21}$$

With the same explanation as in *theorem 4.1*, let $\lambda = 1, 1-\mu = 1 - \frac{4A}{1-4A}$, and we get $\frac{\lambda}{1-\mu} = \frac{1}{1-\frac{4A}{1-4A}} = \frac{1-4A}{1-8A}$ □

4.2.2 More general A: Four Lemmas

A obvious limitation about *theorem 4.1, 4.2* is that when $A \to \frac{1}{2}, \frac{1}{8}$, the upper bound $\to \infty$, in the later part we try to improve this bound for more general A.

We began with the 2-player case to gain some inspirations for more general cases. We use the same settings as the *theorem 4.1*, w.l.o.g let there be J=2k files in total, to minimize the miss probability, a base station won't consider caching files $\{2k+1, 2k+2, ...\}$ as a selfish player. Assume one of the two players caches files $\{i_1, i_2, ..., i_k\}$, the other caches files $\{j_1, j_2, ..., j_k\}$.

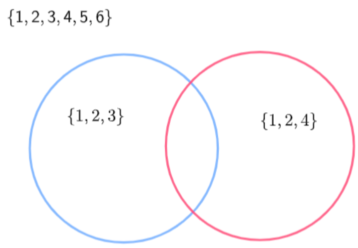


Figure 4: Two Base Stations with Capacity 3

I began with calculating the 2-capacity case to gain some inspirations and I noticed there are some rules between the worst Nash equilibrium and optimal strategy. So I made four guesses and tried to prove them, which are presented as lemma 2,3,4 and 5. Part of enumerations and calculation result is in the appendix A.

Lemma 2. *(Rail Lemma) The possible candidates for Nash equilibrium strategies and optimal strategies are all in the form that first caching the succesive same files and then caching the "complementary" files, i.e. all the strategy profiles which are the candidates of Nash equilibrium and optimal strategy should in the form that:*

- *First have successive same choices, then following complementary choices: one file is only cached by one player.*
- *No jumpings: if* $|\{i_1, i_2, ..., i_k\} \cup \{j_1, j_2, ..., j_k\}| = q$, *then* $\{i_1, i_2, ..., i_k\} \cup \{j_1, j_2, ..., j_k\} = \{1, ..., q\}$ *where* $k \leq q \leq 2k$

Proof. We prove this lemma by contradiction and figures. If the same choices are not successive, as the figure below shows, player 2 can simply choose to deviate to cache file 2 instead of file 4, this "moving the file forward" operation cut down the cost of player 2 and miss probability at the same time because the players share a more popular file now.

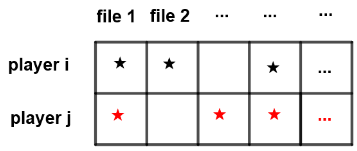


Figure 5: They don't successive same choices

If there are jumpings (files that are not chosen by both players in the middle), both players cut down their cost by a simple "moving forward" operation, and the miss probability decreases at the same time.

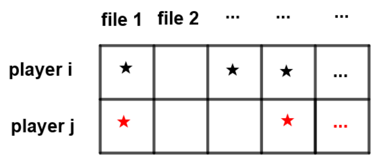


Figure 6: A jumping at location 2

We believe the above figures and explanation have made this problem clear for the readers even if we didn't use formulas and equations. The figures look like rails and that's where this lemma gets the name, a qualified candidate should look like this:

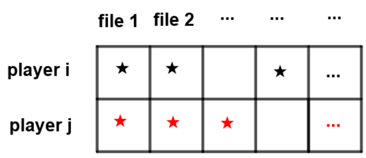


Figure 7: Qualified Candidate

Lemma 2 describes what a possible strategy profile candidate of Nash equilibrium and optimal outcome should look like, this can help us rule out many meaningless cases and will save us much time to explain in the proof of lemma 3, 4 and 5.

□

Lemma 3. *With the basic settings, when* $A \geq 1 - \frac{a_{2k}}{a_1}$ *(best bound), caching complementary files is an optimal and also a Nash equilibrium strategy and this is the only possible combination.*

Remark. There are $\frac{\binom{2k}{k}}{2}$ strategies(consider the symmetry in our model) to cache complementary files but they have the same miss probability :$\frac{1-A}{2-A}$ and function $\frac{1-A}{2-A}$ decreases as A increases.

Proof. This proof consists of two main parts.

1. Caching complementary files is the only Nash equilibrium.

2. Caching complementary files is an optimal strategy profile.

Nash Equilibrium Part
Fix the player i, The cost of the player j is

$$\sum_{k\in\{1,...,2k\}\backslash(\{i_1,i_2,...,i_k\}\cup\{j_1,j_2,...,j_k\})} a_k\frac{1}{2-A}+\sum_{l\in\{i_1,...,i_k\}\backslash\{j_1,...,j_k\}} a_l\frac{1-A}{2-A} \tag{22}$$

especially, when $\{i_1,...,i_k\}$ and $\{j_1,...,j_k\}$ are complementary (i.e. $\{j_1,...,j_k\}=\{1,2,...,2k\}\backslash\{i_1,...,i_k\}$), cost of the second player equals

$$(a_{i_1}+a_{i_2}+...+a_{i_k})\frac{1-A}{2-A} \tag{23}$$

Compare the cost of the second player when caching arbitrary k files and complementary files.

$$\begin{aligned} difference &= (a_{i_1}+a_{i_2}+...+a_{i_k})\frac{1-A}{2-A}-\sum_{l\in\{i_1,...,i_k\}\backslash\{j_1,...,j_k\}} a_l\frac{1-A}{2-A}-\sum_{k\in\{1,...,2k\}\backslash(\{i_1,i_2,...,i_k\}\cup\{j_1,j_2,...,j_k\})} a_k\frac{1}{2-A} \\ &= \sum_{m\in\{i_1,i_2,...,i_k\}\cap\{j_1,j_2,...,j_k\}} a_m\frac{1-A}{2-A}-\sum_{k\in\{1,2,...,2k\}\backslash(\{i_1,...,i_k\}\cup\{j_1,...,j_k\})} a_k\frac{1}{2-A} \end{aligned} \tag{24}$$

According to the range of A, $A\geq 1-\frac{a_{2k}}{a_1}\Leftrightarrow\frac{1-A}{2-A}\leq\frac{a_{2k}}{a_1+a_{2k}}<\frac{a_{2k}}{a_1}$, and $a_{2k}\leq a_l\leq a_1,\forall 1\leq l\leq 2k$, also $|\{i_1,i_2,...,i_k\}\cap\{j_1,j_2,...,j_k\}|=|\{1,2,...,2k\}\backslash(\{i_1,...,i_k\}\cup\{j_1,...,j_k\})|$, we denote it as f. Thus we have

$$\text{difference}<f\times a_1\times\frac{a_{2k}}{a_1}-f\times a_{2k}=0 \tag{25}$$

This means unless the player j caches complementary files, it will definitly deviate.

Optimal Part
We prove this part by induction: Suppose when n=k-1, $A\geq 1-\frac{a_{2k-2}}{a_1}$, caching complementary files is an optimal strategy profile, now let n=k, J=2k, $A\geq 1-\frac{a_{2k}}{a_1}$, we need to prove that caching complementary files is still an optimal strategy profile, we discuss it in two cases:

- Two base stations have different choices
- Two base stations have exactly the same choices, which has to be $\{1,2,...,k\}\times\{1,2,...,k\}$

First case:
Let $\{i_1,i_2,...,i_k\}\times\{j_1,j_2,...,j_k\}$ be an arbitrary optimal strategy profile, if the players have different choices about file i, they must have different choices about another file j, here we assume $i<j$.

Note $A\geq 1-\frac{a_{2k}}{a_1}\geq 1-\frac{a_m}{a_n},\forall m,n\in\{1,2,...,2k\}$, we ignore the file i,j which are both chosen by only one of the base stations, and apply the assumption to the remaining files $\{1,...,2k\}\backslash\{i,j\}$, as players have complementary choices about file i,j, thus $\{i_1,i_2,...,i_k\}\times\{j_1,j_2,...,j_k\}$ is still a complementary strategy profile and we can apply the assumption here.

Second case:
To discuss the case left, we only need to prove that when $A\geq 1-\frac{a_{2k}}{a_1}$, $\{1,2,...,k\}\times\{1,2,...,k\}$ is not optimal, we can easily raise another strategy profile with lower miss probability: $\{1,2,...,k\}\times\{1,2,...,k-1,k+1\}$.

After some simple calculation, miss probability of strategy $\{1,2,...,k\}\times\{1,2,...,k-1,k+1\}$ equals

$$(a_{k+1}+...+a_{2k})+\frac{(1-A)a_k-(3-2A)a_{k+1}}{2-A}$$

We can find that

$$(a_{k+1}+...+a_{2k})+\frac{(1-A)a_k-(3-2A)a_{k+1}}{2-A}<a_{k+1}+...+a_{2k}$$

□

Lemma 4. *With basic settings, when $A \leq 1 - \frac{a_{k+1}}{a_k}$ (best bound), caching $\{1, 2, ..., k\} \times \{1, 2, ..., k\}$ is a Nash equilibrium and also the optimal strategy, this is also the only possible combination.*

Remark. note: $x^{-\gamma} + (2k + 1 - x)^{-\gamma}$ decreases as x increases where $\gamma > 0$

Proof. We also complete this proof in two main parts: Nash equilibrium part and optimal part.

Let $\{i_1, i_2, ..., i_k\} \times \{j_1, j_2, ..., j_k\}$ be an arbitrary Nash equilibrium. By *lemma 2*, we know what the startegy profile exactly looks like: successive same choices and then complementary choices. If they have different choices on file i, there must be another file j about which they make different choices.

Assume their first different choice is on file j, if $1 \leq j \leq k$, there must be some i satisfying $k + 1 \leq i \leq 2k$ about which they also make different choices because they have the same capacity, in figure 8, file j is on the second position and file i is on the fourth position.

Nash Equilibrium Part

We suggest the player to cache file j instead of file i, now we examine the change of the cost of the player to see if our suggestion will be adopted:

$$a_i \frac{1}{2 - A} - a_j \frac{1 - A}{2 - A} \tag{26}$$

where $1 - A \geq \frac{a_{k+1}}{a_k}$, so the selfish player will absolutely adopt this wise suggestion, i.e. deviate, this proof also applies when player i caches $\{1, 2, ..., k\}$ thus we know $\{1, 2, ..., k\} \times \{1, 2, ..., k\}$ is the only Nash equilibrium while $A \leq 1 - \frac{a_{k+1}}{a_k}$.

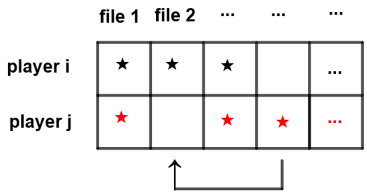


Figure 8: Wise Suggestion
file j is on the second position and file i is on the fourth position

Optimal Part

We calculate the miss probability of an arbitrary optimal strategy profile, assume they both have chosen to cache the first l $(0 \leq l \leq k)$ files and then the complementary files follow.

$$a_{2k-l+1} + ... + a_{2k} + (a_{l+1} + ... + a_{2k-l}) \frac{1 - A}{2 - A}$$

$A \leq 1 - \frac{a_{k+1}}{a_k} \Leftrightarrow \frac{1-A}{2-A} \geq \frac{a_{k+1}}{a_k + a_{k+1}}$. As l gets larger the above fomular will get smaller because $a_i - a_j \frac{1-A}{2-A} \leq a_{k+1} - (a_k + a_{k+1}) \frac{1-A}{2-A} \leq 0, \forall i, j \in \{1, 2, ..., 2k\}$ and apply the remark above.

□

Another observation about during the calculation is that some of the combinations can not exist, e.g. Combinations such as $\{1, 2, 3\} \times \{1, 2, 3\}_{player1}, \{1, 2, 4\} \times \{1, 3, 5\}_{player2}$ as Nash equilibrium and optimal strategy (or optimal strategy and Nash equilibrium) does not exist,$\{1, 2, 3\} \times \{1, 2, 3\}_{player1}, \{1, 3, 5\} \times \{2, 4, 6\}_{player2}$ neither. From these observation, *lemma 5* is construed.

Lemma 5. *The completely same and completely complementary case strategy profile is impossible to "coexist": One is Nash equilibrium while the other is optimal, here we assume the capacity to be strictly larger than 1.*

Proof. Assume $l = k$ is the optimal strategy and $l = 0$ is the Nash equilibrium, apply necessary optimal condition, miss probability of $l = k$ case is less than $l = k - 1$-case:

$$\begin{aligned} a_{k+1} + ... + a_{2k} &\leq a_{k+2} + ... + a_{2k} + (a_k + a_{k+1}) \frac{1 - A}{2 - A} \\ A &\leq 1 - \frac{a_{k+1}}{a_k} \end{aligned} \tag{27}$$

Then apply necessary Nash condition, player j won't deviate to cache file 1 instead of file j_k (there has to be a player who doesn't cache file 1, we can just assume player j w.l.o.g.):

$$\begin{aligned} a_1 \frac{1-A}{2-A} - a_{j_k} \frac{1}{2-A} \leq 0 \\ A \geq 1 - \frac{a_{j_k}}{a_1} \geq 1 - \frac{a_{k+1}}{a_1} \end{aligned} \tag{28}$$

The intersection of these two intervals is empty, we get a contradiction.

□

5 Unfinished Work

There are many limitations about the analysis methods adopted above, such as you can't generalize 2-player case to N players case by simply look at them in pairs, because the benefit by caching a file now is influenced by the third player, but you can still find that the jumping is not possible for N players case.

As for the files which are not cached by the third player, it's the same in the rail lemma, but for the files chosen by the third player, the third player will dilute the benefit.

It's kind of complicated to analyze using the same method as in the lemma 1, and to get an upper bound of the 2-players case when $1 - \frac{a_{k+1}}{a_k} \leq A \leq 1 - \frac{a_1}{a_{2k}}$, it's just a beginning, here are some ideas may work in the future study:

- Discuss the range of the A, when $l = 0, 1, ..., k$ is the optimal strategy and what's the worst Nash equilibrium

$$1 - \frac{a_{2k-l}}{a_{l+1}} \leq A \leq 1 - \frac{a_{2k+1-l}}{a_l}$$

  $A \geq 1 - \frac{a_{2k}}{a_1}, A \leq 1 - \frac{a_{k+1}}{a_k}$ have been discussed.

- Assume $i \times j$ is the worst Nash equilibrium, i.e. $min\{i, j\} - l$ reaches the maximum or minimum.
  - For arbitrary fixed strategy j, the set of strategies which player i most probably deviate to consist of strategies cache consecutive files and then complementary with player i, but note here the lemma 1 doesn't apply because for now we don't have to consider bilateral deviation.
- As for the derivative method, a better bound should be developed for $\sum_{i=1}^{N} c_i(s_i^*, s_{-i})$ and $\sum_{i=1}^{N} c_i(s)$

6 Blocked roads: failed attempts on equivalent game method

Here we just list the main results below, please see the *appendix B* for more details.

6.1 Congestion Game: Failed because we use the wrong social cost function

- Build the corresponding congestion game model.
- Construe examples with bad PoA and find out why.
- Tried smoothness and m,n version $(\lambda - \mu)$-smoothness.
- The indicator function in cost functions makes the method not applicable.

6.2 Market Sharing Game-A Special Case of Congestion game

- Market sharing game is a valid utility system and a valid utility system has an upper bound 2 on PoA.
- Failed because our game doesn't satisfy some critical condition for a valid utility system.

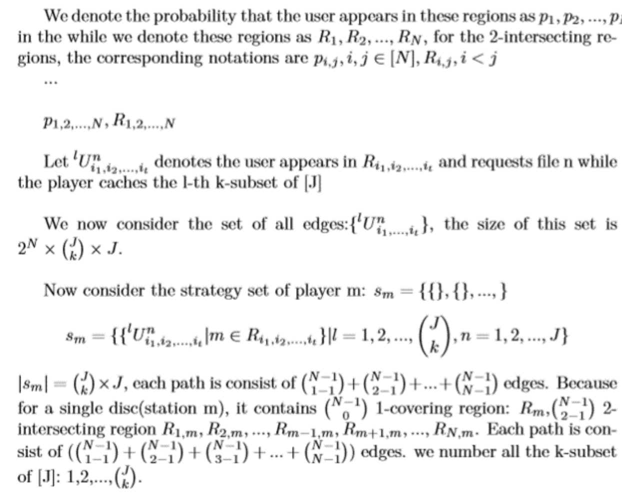
6.3 Wrong way to construe derivative inequality

- Didn't use the terms with intersecting area A but obtained an upper bound depend on the number of files instead.

7 Gains and Summary

Apart from the results, methods applied in this project are also worth mentioning. They are super useful and extremely important for conducting science researches and probably any other problems.

- Importance of abstraction: when building the equivalent game model with congestion game model, if you try to use the "easy" corresponding of: base stations $\leftrightarrow$ players, files $\leftrightarrow$ resources, you probably will get an "ugly" notation such as ${}^{l}U_i^n$ and complicated model like in the picture below.



We denote the probability that the user appears in these regions as $p_1, p_2, ..., p_N$, in the while we denote these regions as $R_1, R_2, ..., R_N$, for the 2-intersecting regions, the corresponding notations are $p_{i,j}, i, j \in [N], R_{i,j}, i < j$

...

$p_{1,2,...,N}, R_{1,2,...,N}$

Let ${}^{l}U^n_{i_1,i_2,...,i_t}$ denotes the user appears in $R_{i_1,i_2,...,i_t}$ and requests file n while the player caches the l-th k-subset of [J]

We now consider the set of all edges:$\{{}^{l}U^n_{i_1,...,i_t}\}$, the size of this set is $2^N \times \binom{J}{k} \times J$.

Now consider the strategy set of player m: $s_m = \{\{\}, \{\}, ..., \}$

$$s_m = \{\{{}^{l}U^n_{i_1,i_2,...,i_t} | m \in R_{i_1,i_2,...,i_t}\} | l = 1, 2, ..., \binom{J}{k}, n = 1, 2, ..., J\}$$

$|s_m| = \binom{J}{k} \times J$, each path is consist of $\binom{N-1}{1-1} + \binom{N-1}{2-1} + ... + \binom{N-1}{N-1}$ edges. Because for a single disc(station m), it contains $\binom{N-1}{0}$ 1-covering region: R_m,$\binom{N-1}{2-1}$ 2-intersecting region $R_{1,m}, R_{2,m}, ..., R_{m-1,m}, R_{m+1,m}, ..., R_{N,m}$. Each path is consist of $(\binom{N-1}{1-1} + \binom{N-1}{2-1} + \binom{N-1}{3-1} + ... + \binom{N-1}{N-1})$ edges. we number all the k-subset of [J]: 1,2,...,$\binom{J}{k}$.

Figure 9: Ugly Model

  But once you mix incidences into the model, all become clear and elegant at once like in appendix B. Take the most obvious choice now might make your life much harder in the later.

- You should always check your results with simple examples first.

- When you find it difficult to analyse a problem directly, try to disentangle it. Start with the simple cases and be patient with them, they might seem too trivial to stir any inspirations at first, but that's probably you have not tried enough examples and observe enough, the trick is: nobody wants to do enumerations so think about how to make your life easier, what is repeated during this brute force method. Just like you define a function to solve the similar problems when coding, you can also make a guess, try to prove it, and you get a lemma to do the repeated analysis work for you. This is an essential trick to obtain the *lemma 2-5* in this report.
  - One last thing about the last summary: be prepared that the enumeration work can be a little more than you expected, but it's a good way to force you to keep your draft clear and ordered.

- We can always use assumptions to make things easier, i.e make some concessions. This one applies to the *theorem 4.1, 4.2*, then you can move ahead little by little.

- Learning from bad examples: opposite of what you want to prove, such as construing examples with bad PoA and find out what causes this? What makes it deviate from the optimal outcome, this method might not be so useful in our results but it make you understand some models better, in our case, *pigou's game.*.

- Make a circle to avoid boundaries. See *Theorem 4.2*.

- Accumulation is the base of everthing above.
- Maybe (Probably) you will take many wrong paths before you find a hopeful one, so keep healthy and happy to be prepared for that.

8 reference

References

[1] Konstantin Avrachenkov, Jasper Goseling, and Berksan Serbetci. 2017. A Low-Complexity Approach to Distributed Cooperative Caching with Geographic Constraints. Proc. ACM Meas. Anal. Comput. Syst. 1, 1, Article 27 (June 2017), 24 pages.

[2] Tim Roughgarden. Lectures on Algorithm Game Theory

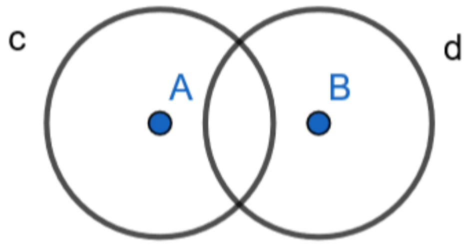
[3] A. Vetta, "Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions," The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings., Vancouver, BC, 2002, pp. 416-425.

A Enumeration Work

Most tiring enumeration work is on my notebooks and I will do anything but typing them again with latex.

A.1 Two Player Capacity Two-case

We begin with the two player case.



	base station A	base station B
cost	$\frac{a_2}{2-s}$	$\frac{a_2}{2-s}$
Miss Prob.		a_2

Table 1: strategy one:A caches 1, B caches 1

	base station A	base station B
cost	$\frac{a_2(1-s)}{2-s}$	$\frac{a_1(1-s)}{2-s}$
Miss Prob.		$\frac{(a_1+a_2)(1-s)}{2-s}$

Table 2: strategy two: A caches 1, B caches 2

So there are four cases we need to discuss:
Following are the inequalities due to the definitioin of optimal outcome and Nash equilibrium.

	Optimal Outcome	Nash Equilibrium
case 1	strategy 1	strategy 1
case 2	strategy 1	strategy 2
case 3	strategy 2	strategy 1
case 4	strategy 2	strategy 2

Table 3: Possible Combinations

$$\begin{cases} \frac{a_2}{2-s} \leq \frac{a_1(1-s)}{2-s} \\ a_2 \leq \frac{a_1(1-s)}{2-s} + \frac{a_2(1-s)}{2-s} \end{cases} \tag{29}$$

$$\begin{cases} \frac{a_2}{2-s} \leq \frac{a_1(1-s)}{2-s} \\ a_2 \leq \frac{a_1(1-s)}{2-s} + \frac{a_2(1-s)}{2-s} \end{cases} \tag{30}$$

$$\begin{cases} \frac{a_2}{2-s} \leq \frac{a_1(1-s)}{2-s} \\ a_2 \leq \frac{a_1(1-s)}{2-s} + \frac{a_2(1-s)}{2-s} \end{cases} \tag{31}$$

$$\begin{cases} \frac{a_2}{2-s} \leq \frac{a_1(1-s)}{2-s} \\ a_2 \leq \frac{a_1(1-s)}{2-s} + \frac{a_2(1-s)}{2-s} \end{cases} \tag{32}$$

After solving the inequalities above, we obtain the corresponding interval of s in each case:

case 1	$s \leq \frac{a_1-a_2}{a_1}$
case 2	$s = \frac{a_1-a_2}{a_1}$
case 3	$s = \frac{a_1-a_2}{a_1}$
case 4	$1 \geq s \geq \frac{a_1-a_2}{a_1}$

Table 4: Corresponding ranges of s

A.2 Three Players Capacity One-case:

Then we move to the case of three base stations and all of them having the same distance, here we set the radius to 1 and denote the distance by d, i.e.$|AB| = |AC| = |BC| = d$. Denote the area of $region_{ADHN}$ by x, area of $region_{EGHD}$ by y, area of $region_{GKH}$ by z.

$$\begin{cases} x + 2y + z = \frac{\pi}{6} \\ y + z = \frac{2arccos\frac{d}{2} - d\sqrt{1-\frac{d^2}{4}}}{2} \\ 3x + 3y + z = \frac{\sqrt{3}d^2}{4} \end{cases} \tag{33}$$

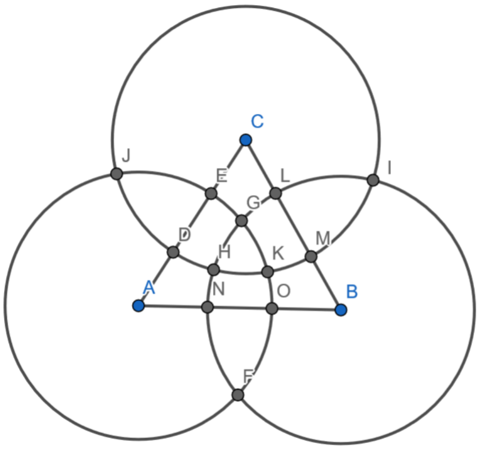
By solving the equations above, we get the area ratio of three different types of region: 1-covering region, 2-intersecting region, 3-intersecting region.

$$\begin{cases} x = -\frac{\pi}{3} + arccos\frac{d}{2} - \frac{d\sqrt{1-\frac{d^2}{4}}}{2} + \frac{\sqrt{3}d^2}{4} \\ y = -2arccos\frac{d}{2} + \frac{d\sqrt{1-\frac{d^2}{4}}}{2} - \frac{\sqrt{3}d^2}{4} + \frac{\pi}{2} \\ z = \frac{\sqrt{3}d^2}{4} - \frac{\pi}{2} + 3arccos\frac{d}{2} - \frac{3d\sqrt{1-\frac{d^2}{4}}}{2} \end{cases} \tag{34}$$

$$\begin{cases} p_1 = p_2 = p_3 = \frac{\pi-(3z+4y)}{s_{all}} \\ p_{1,2} = p_{2,3} = p_{1,3} = \frac{2y+z}{s_{all}} \\ p_{1,2,3} = \frac{z}{s_{all}} \end{cases} \tag{35}$$

$$S_{all} = 3\pi - 5z - 6y$$

But the notations above are not good for gaining an intuition about relationships between the intersecting area and PoA.



So we just denote the 1-covering areas as O, 2-covering area as W, 3-covering area as T.

$$O = \pi - (3z + 4y)$$

$$W = 2y + z$$

$$T = z$$

We know that O+2W+T=π, $S_{all} = 3O + 3W + T$, with capacity 1, 3 files, there are 3 optimal strategy candidates and 3 Nash equilibrium candidates. Thus there are 9 cases we need to discuss.

	Base Station 1	Base Station 2	Base Station 3
strategy 1	file 1	file 1	file 1
strategy 2	file 1	file 1	file 2
strategy 3	file 1	file 2	file 3

Table 5: Nash Equilibrium and Optimal Candidates

	Optimal Outcome	Nash Equilibrium
case 1	1,1,1	1,1,1
case 2	1,1,1	1,1,2
case 3	1,1,1	1,2,3
case 4	1,1,2	1,1,1
case 5	1,1,2	1,1,2
case 6	1,1,2	1,2,3
case 7	1,2,3	1,1,1
case 8	1,2,3	1,1,2
case 9	1,2,3	1,2,3

Table 6: Possible Combinations

Case 1

	base station A:1	base station B:1	base station C:1
cost	$\frac{(a_2+a_3)\pi}{S_{all}}$	$\frac{(a_2+a_3)\pi}{S_{all}}$	$\frac{(a_2+a_3)\pi}{S_{all}}$
Miss Prob.		a_2+a_3	

Table 7: strategy one: 1,1,1

	base station A:1	base station B:1	base station C:2
cost	$\frac{a_2(O+W)}{S_{all}}+\frac{a_3\pi}{S_{all}}$	$\frac{a_2(O+W)}{S_{all}}+\frac{a_3\pi}{S_{all}}$	$\frac{a_1O}{S_{all}}+\frac{a_3\pi}{S_{all}}$
Miss Prob.		$\frac{a_2(2O+W)}{S_{all}}+\frac{a_1O}{S_{all}}+a_3$	

Table 8: strategy two: 1,1,2

	base station A:1	base station B:2	base station C:3
cost	$\frac{(a_2+a_3)(O+W)}{S_{all}}$	$\frac{(a_1+a_3)(O+W)}{S_{all}}$	$\frac{(a_1+a_2)(O+W)}{S_{all}}$
Miss Prob.		$\frac{(a_1+a_2+a_3)(2O+W)}{S_{all}}$	

Table 9: strategy three: 1,2,3

	base station A:1	base station B:1	base station C:3
cost	$\frac{a_3(O+W)}{S_{all}}+\frac{a_2\pi}{S_{all}}$	$\frac{a_3(O+W)}{S_{all}}+\frac{a_2\pi}{S_{all}}$	$\frac{a_1O}{S_{all}}+\frac{a_2\pi}{S_{all}}$
Miss Prob.		$\frac{a_3(2O+W)}{S_{all}}+\frac{a_1O}{S_{all}}+a_2$	

Table 10: strategy four: 1,1,3

	base station A:1	base station B:2	base station C:2
cost	$\frac{a_2O}{S_{all}}+\frac{a_3\pi}{S_{all}}$	$\frac{a_1(O+W)}{S_{all}}+\frac{a_3\pi}{S_{all}}$	$\frac{a_1(O+W)}{S_{all}}+\frac{a_3\pi}{S_{all}}$
Miss Prob.		$\frac{a_1(2O+W)}{S_{all}}+\frac{a_2O}{S_{all}}+a_3$	

Table 11: strategy five: 1,2,2

$$\text{case 1}=\begin{cases} a_2+a_3\le \frac{a_2(2O+W)}{S_{all}}+\frac{a_1O}{S_{all}}+a_3(\text{cost strategy 1 is less than strategy 2})\\ a_2+a_3\le \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}}(\text{cost strategy 1 is less than strategy 3})\\ \frac{(a_2+a_3)\pi}{S_{all}}\le \frac{a_1O}{S_{all}}+\frac{a_3\pi}{S_{all}}(\text{1 won't deviate to 2(1,1,1}\to\text{1,1,2))}\\ \frac{(a_2+a_3)\pi}{S_{all}}\le \frac{a_1O}{S_{all}}+\frac{a_2\pi}{S_{all}}(\text{1 won't deviate to 3(1,1,1}\to\text{1,1,3))} \end{cases} \tag{36}$$

The first two inequalities are from the definition of optimal outcome, and the third and fourth inequalities are from the definition of Nash equilibrium. Note that we only need the third inequality because the RHS of the fourth inequality is greater than the fourth inequality.

And note the relationships below:

$$p_1=p_2=p_3=\frac{O}{S_{all}}$$

$$p_{1,2}=p_{1,3}=p_{2,3}=\frac{W}{S_{all}}$$

$$p_{1,2,3}=\frac{T}{S_{all}}$$

We solve the group of inequalities in (8) first. After simplifying it:

$$(1)\Leftrightarrow O\ge\frac{\pi a_2}{a_1}$$

$$(2)\Leftrightarrow 2O+W\ge\frac{\pi(a_2+a_3)}{a_1}$$

$$(3) \Leftrightarrow O \geq \frac{\pi a_2}{a_1}$$

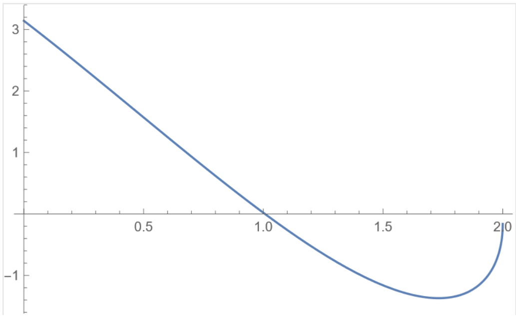


Figure 10: $f(d) = 3z(d) + 4y(d)$, to get $O \geq \frac{a_2\pi}{a_1} \Leftrightarrow 3z + 4y \leq \frac{\pi(a_1-a_2)}{a_1}$

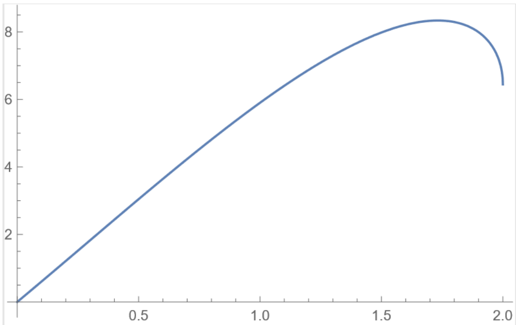


Figure 11: $f(d) = 2O + W$, to get $2O + W \geq \frac{\pi(a_2+a_3)}{a_1}$

Thus in case one, we get the interval of d in the form of [some expression depends on $\gamma, +\infty$)
case 2

$$\text{case 2}=\begin{cases} a_2 + a_3 \leq \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 (\text{cost of strategy 1 is less than strategy 2}) \\ a_2 + a_3 \leq \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} (\text{cost of strategy 1 is less than strategy 3}) \\ \frac{a_1 O}{S_{all}} + \frac{a_3\pi}{S_{all}} \leq \frac{(a_2+a_3)\pi}{S_{all}} (\text{2 won't deviate to 1(i.e. 1,1,2}\rightarrow\text{ 1,1,1))} \\ \frac{a_2(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} \leq \frac{(a_1+a_2)(O+W)}{S_{all}} (\text{1 won't deviate to 3(i.e. 1,1,2}\rightarrow\text{1,3,2))} \end{cases} \tag{37}$$

After simplification:

$$(1) \Leftrightarrow O \geq \frac{\pi a_2}{a_1}$$

$$(2) \Leftrightarrow 2O + W \geq \frac{\pi(a_2 + a_3)}{a_1}$$

$$(3) \Leftrightarrow O \leq \frac{\pi a_2}{a_1}$$

$$(4) \Leftrightarrow O + W \geq \frac{a_3\pi}{a_1}$$

We obtain that $O = \frac{\pi a_2}{a_1}$ from the four inequalities above.
case 3

$$\text{case 3}=\begin{cases} a_2 + a_3 \leq \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 (\text{cost of strategy 1 is less than strategy 2}) \\ a_2 + a_3 \leq \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} (\text{cost of strategy 1 is less than strategy 3}) \\ \frac{(a_1+a_3)(O+W)}{S_{all}} \leq \frac{a_3(O+W)}{S_{all}} + \frac{a_2\pi}{S_{all}} (\text{2 won't deviate to 1(i.e. 1,2,3}\rightarrow\text{ 1,1,3))} \\ \frac{(a_1+a_2)(O+W)}{S_{all}} \leq \frac{a_2(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} (\text{3 won't deviate to 1(i.e. 1,2,3}\rightarrow\text{1,2,1))} \\ \frac{(a_1+a_2)(O+W)}{S_{all}} \leq \frac{a_1(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} (\text{3 won't deviate to 2(i.e. 1,2,3}\rightarrow\text{1,2,2))} \end{cases} \tag{38}$$

After simplification:

$$(1) \Leftrightarrow O \geq \frac{\pi a_2}{a_1}$$

$$(2) \Leftrightarrow 2O + W \geq \frac{\pi(a_2 + a_3)}{a_1}$$

$$(3) \Leftrightarrow O + W \leq \frac{a_2\pi}{a_1}$$

$$(4) \Leftrightarrow O + W \leq \frac{a_3\pi}{a_1}$$

$$(5) \Leftrightarrow O + W \leq \frac{a_3\pi}{a_2}$$

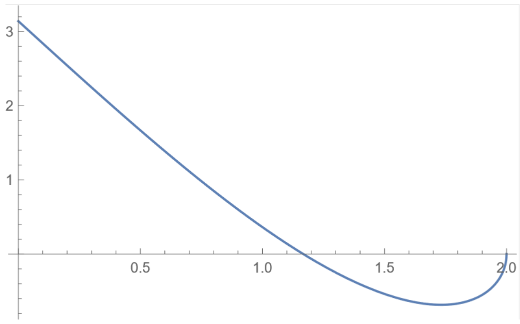


Figure 12: $f(d) = 2y + 2z$, to get $O + W \leq \frac{a_3\pi}{a_1} \Leftrightarrow 2z + 2y \geq \frac{\pi(a_1 - a_3)}{a_1}$

We may get an interval of d in the form of $[L, U], L > 0, U < 2$ and they depend on γ in case 3.
case 4

$$\text{case 4=}\begin{cases} \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 \leq a_2 + a_3 (\text{cost of strategy 2 is less than strategy 1}) \\ \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 \leq \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} (\text{cost of strategy 2 is less than strategy 3}) \\ \frac{(a_2+a_3)\pi}{S_{all}} \leq \frac{a_1 O}{S_{all}} + \frac{a_3\pi}{S_{all}} (\text{1 won't deviate to 2(1,1,1}\to\text{1,1,2))} \\ \frac{(a_2+a_3)\pi}{S_{all}} \leq \frac{a_1 O}{S_{all}} + \frac{a_2\pi}{S_{all}} (\text{1 won't deviate to 3(1,1,1}\to\text{1,1,3))} \end{cases} \quad (39)$$

After simplification:

$$(1) \Leftrightarrow O \leq \frac{a_2\pi}{a_1}$$

$$(2) \Leftrightarrow O + W \geq \frac{a_3\pi}{a_1}$$

$$(3) \Leftrightarrow O \geq \frac{a_2\pi}{a_1}$$

We get that $O = \frac{a_2\pi}{a_1}$ in case 4.
case 5

$$\text{case 5=}\begin{cases} \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 \leq a_2 + a_3 (\text{cost of strategy 2 is less than strategy 1}) \\ \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 \leq \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} (\text{cost of strategy 2 is less than strategy 3}) \\ \frac{a_1 O}{S_{all}} + \frac{a_3\pi}{S_{all}} \leq \frac{(a_2+a_3)\pi}{S_{all}} (\text{2 won't deviate to 1(i.e. 1,1,2}\to\text{ 1,1,1))} \\ \frac{a_2(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} \leq \frac{(a_1+a_2)(O+W)}{S_{all}} (\text{1 won't deviate to 3(i.e. 1,1,2}\to\text{1,3,2))} \end{cases} \quad (40)$$

After simplification:

$$(1) = (3) \Leftrightarrow O \leq \frac{a_2\pi}{a_1}$$

$$(2) = (4) \Leftrightarrow O + W \geq \frac{a_3 \pi}{a_1}$$

We may get an interval of d in the form of $[L, U], L > 0, U < 2$ and they depend on γ in case 5.
case 6

$$\text{case 6}=\begin{cases} \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 \leq a_2 + a_3 (\text{cost of strategy 2 is less than strategy 1}) \\ \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 \leq \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} (\text{cost of strategy 2 is less than strategy 3}) \\ \frac{(a_1+a_3)(O+W)}{S_{all}} \leq \frac{a_3(O+W)}{S_{all}} + \frac{a_2\pi}{S_{all}} (\text{2 won't deviate to 1(i.e. 1,2,3} \rightarrow \text{1,1,3))}) \\ \frac{(a_1+a_2)(O+W)}{S_{all}} \leq \frac{a_2(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} (\text{3 won't deviate to 1(i.e. 1,2,3} \rightarrow \text{1,2,1))}) \\ \frac{(a_1+a_2)(O+W)}{S_{all}} \leq \frac{a_1(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} (\text{3 won't deviate to 2(i.e. 1,2,3} \rightarrow \text{1,2,2))}) \end{cases} \quad (41)$$

After simplification:

$$(1) \Leftrightarrow O \leq \frac{a_2 \pi}{a_1}$$

$$(2) \Leftrightarrow O + W \geq \frac{a_3 \pi}{a_1}$$

$$(3) \Leftrightarrow O + W \leq \frac{a_2 \pi}{a_1}$$

$$(4) \Leftrightarrow O + W \leq \frac{a_3 \pi}{a_1}$$

$$(5) \Leftrightarrow O + W \leq \frac{a_3 \pi}{a_2}$$

We get $O + W = \frac{a_3 \pi}{a_1}$
case 7

$$\text{case 7}=\begin{cases} \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} \leq a_2 + a_3 (\text{cost of strategy 3 is less than strategy 1}) \\ \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} \leq \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 (\text{cost of strategy 3 is less than strategy 2}) \\ \frac{(a_2+a_3)\pi}{S_{all}} \leq \frac{a_1 O}{S_{all}} + \frac{a_3\pi}{S_{all}} (\text{1 won't deviate to 2(1,1,1} \rightarrow \text{1,1,2))}) \\ \frac{(a_2+a_3)\pi}{S_{all}} \leq \frac{a_1 O}{S_{all}} + \frac{a_2\pi}{S_{all}} (\text{1 won't deviate to 3(1,1,1} \rightarrow \text{1,1,3))}) \end{cases} \quad (42)$$

After simplification:

$$(1) \Leftrightarrow O \leq \frac{\pi a_2}{a_1}$$

$$(2) \Leftrightarrow O + W \leq \frac{a_3 \pi}{a_1}$$

$$(3) \Leftrightarrow O \geq \frac{a_2 \pi}{a_1}$$

Case 7 does not exist.
case 8

$$\text{case 8}=\begin{cases} \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} \leq a_2 + a_3 (\text{cost of strategy 3 is less than strategy 1}) \\ \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} \leq \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3 (\text{cost of strategy 3 is less than strategy 2}) \\ \frac{a_1 O}{S_{all}} + \frac{a_3\pi}{S_{all}} \leq \frac{(a_2+a_3)\pi}{S_{all}} (\text{2 won't deviate to 1(i.e. 1,1,2} \rightarrow \text{1,1,1))}) \\ \frac{a_2(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}} \leq \frac{(a_1+a_2)(O+W)}{S_{all}} (\text{1 won't deviate to 3(i.e. 1,1,2} \rightarrow \text{1,3,2))}) \end{cases} \quad (43)$$

After simplification:

$$(1) \Leftrightarrow O \leq \frac{\pi a_2}{a_1}$$

$$(2) \Leftrightarrow O + W \leq \frac{a_3 \pi}{a_1}$$

$$(3) \Leftrightarrow O \leq \frac{a_2\pi}{a_1}$$

$$(4) \Leftrightarrow O + W \geq \frac{a_3\pi}{a_1}$$

Solve the above inequations, we get $O + W = \frac{a_3\pi}{a_1}$ in case 8.

case 9

$$\text{case 9=}\begin{cases} \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} \leq a_2 + a_3(\text{cost of strategy 3 is less than strategy 1}) \\ \frac{(a_1+a_2+a_3)(2O+W)}{S_{all}} \leq \frac{a_2(2O+W)}{S_{all}} + \frac{a_1 O}{S_{all}} + a_3(\text{cost of strategy 3 is less than strategy 2}) \\ \frac{(a_1+a_3)(O+W)}{S_{all}} \leq \frac{a_3(O+W)}{S_{all}} + \frac{a_2\pi}{S_{all}}(\text{2 won't deviate to 1(i.e. 1,2,3}\rightarrow\text{ 1,1,3))} \\ \frac{(a_1+a_2)(O+W)}{S_{all}} \leq \frac{a_2(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}}(\text{3 won't deviate to 1(i.e. 1,2,3}\rightarrow\text{1,2,1))} \\ \frac{(a_1+a_2)(O+W)}{S_{all}} \leq \frac{a_1(O+W)}{S_{all}} + \frac{a_3\pi}{S_{all}}(\text{3 won't deviate to 2(i.e. 1,2,3}\rightarrow\text{1,2,2))} \end{cases} \quad (44)$$

After simplification:

$$(1) \Leftrightarrow O \leq \frac{\pi a_2}{a_1}$$

$$(2) \Leftrightarrow O + W \leq \frac{a_3\pi}{a_1}$$

$$(3) \Leftrightarrow O + W \leq \frac{a_2\pi}{a_1}$$

$$(4) \Leftrightarrow O + W \leq \frac{a_3\pi}{a_1}$$

$$(5) \Leftrightarrow O + W \leq \frac{a_3\pi}{a_2}$$

Thus we get the interval of d in the form of [0, some expression depends on γ].

Results when γ=1

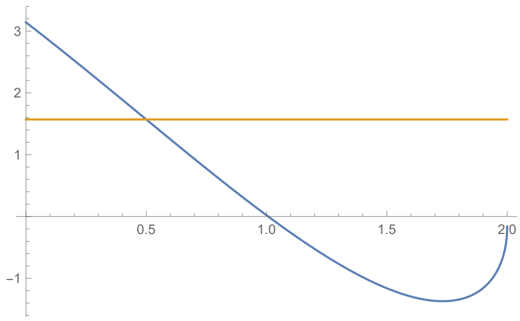


Figure 13: $f(d) = 2O + W$, to get $2O + W = \frac{\pi(a_2+a_3)}{a_1}$

We denote the intersecting points by x_1, x_2, x_3, and the relationship among them is $x_1 > x_2 > x_3$

From the table 12 we observed interesting symmetry in the result.

B Failed Attempts

B.1 Failed Equivalent Congestion Game Model

B.1.1 Technical preliminaries

We try to translate the model in the paper to the game theory language in the lectures. It's easy to write the cost function of a player if we adopt the obvious thought like collaborating or not, but it makes the further steps extremely strenuous: to generalize the model to more than two base stations and two files and so on.

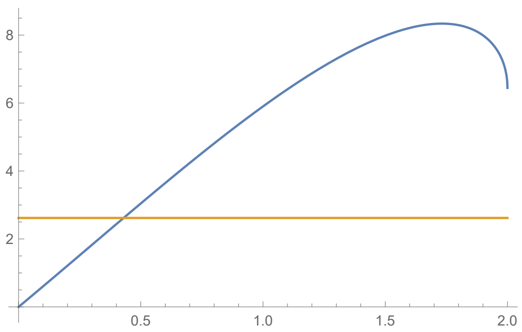


Figure 14: $f(d) = 2O + W$, to get $2O + W = \frac{\pi(a_2+a_3)}{a_1}$

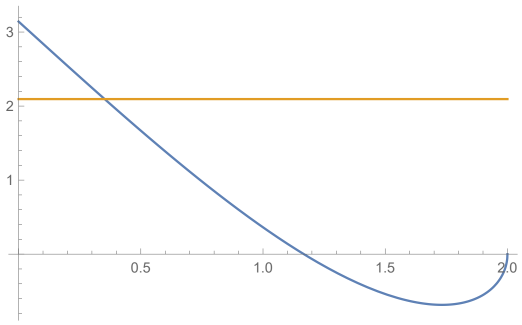


Figure 15: $f(d) = 2y + 2z = \frac{\pi(a_1-a_3)}{a_1}$, to get $O + W \leq \frac{a_3\pi}{a_1} \Leftrightarrow 2z + 2y = \frac{\pi(a_1-a_3)}{a_1}$

	d
case 1	$[x_1, +\infty]$
case 2	$d = x_1$
case 3	does not exist
case 4	$d = x_1$
case 5	$[x_3, x_1]$
case 6	$d = x_3$
case 7	does not exist
case 8	$d = x_3$
case 9	$[0, x_3]$

Table 12: Results when $\gamma = 1$

In selfish routing game, P is the set of paths, p denotes an arbitrary path, f : how traffic splits over paths, f_e: the number of players (base stations) who choose path e, $c(f)$: the total travel time incurred by traffic. Here are some important equations below

$$c_p(f) = \sum_{e \in p} c_e(f_e)$$

$$c(f) = \sum_{p \in P} f_p c_p(f)$$

$$c(f) = \sum_{e \in E} f_e c_e(f_e)$$

Now focus on the first equation, which describes a relationship between objective function (cost function in the paper) and parts which it is constituted of.

Forget about all the vertices, edges and paths, we try to disentangle the objective function through its summation form. To do this, we need to consider the path in a more general sense, the edges here we defined are events, and these two edges constitute a path where the base station j chooses to cache the file i- an element in the strategy set of base station j, we will describe the model in attempt two in detail in the following part.

change from

$$c(f) = \sum_{p \in P} f_p c_p(f)$$

to

$$c(f) = \sum_{i=1}^{k} f_{p_i} c_{p_i}(f)$$

$$= \sum_{i=1}^{k} f_{p_i} \sum_{e \in p_i} c_e(f_e)$$

The last equation holds because the fact : $\sum_{e \in p_i} f_e = f_{p_i}$. Then we'd like to generalize our conclusion to N base stations (players) with capacity k, J files. With N base stations, we can define $\binom{N}{1}+\binom{N}{2}+\binom{N}{3}+...+\binom{N}{N} = 2^N$ intersecting region, N sole-covering region, $\binom{N}{2}$ 2-intersecting region and so on. Instead of focusing on what base station has cached, we can make our work much easier by paying more attention to the file that requested by the user but not in the storation of the base station.

Now we consider the path set(strategy set) of base station m,$m \in [N]$

$$s_m = \{A_s^m | s \subseteq \{1, 2, ..., J\}, |s| = k\}$$

where $A_s^m = \{U_T^n | m \in T, n \notin s\}$, and we list some trivial facts below:

1. $|s_m| = \binom{J}{k}$
2. $|A_s^m| = (J-k) \times [\binom{N-1}{1-1} + \binom{N-1}{2-1} + ... + \binom{N-1}{N-1}] = (J-k) \times 2^{N-1}$

Denote the probability that the user appears within the region covered only by base stations in T ($T \subset \{1, 2, ..., N\}$)as p_T. To obtain the general expression of $c_e(f_e)$, firstly we look at some simple while specific examples.

B.1.2 2 players with capacity 1, 2 files

a: the user chooses file 1 in region 1 (only base station 1 covering)
b: the user chooses file 1 in region 2 (intersecting region of base station 1 and 2)
c: the user chooses file 1 in region 3 (only base station 2 covering)
d: the user chooses file 2 in region 1
e: the user chooses file 2 in region 2
f: the user chooses file 2 in region 3

$$s_1 = \{\{a, b\}, \{d, e\}\}$$

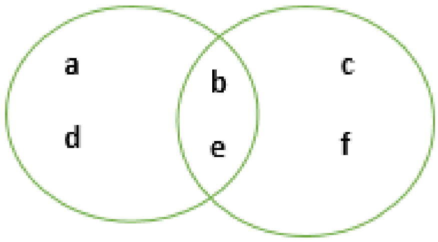


Figure 16: figure title

$$s_2 = \{\{b, c\}, \{e, f\}\}$$

$\{a, b\}$ denotes base station 1 chooses to cache file 1
$\{d, e\}$ denotes base station 1 chooses to cache file 2
$\{b, c\}$ denotes base station 2 chooses to cache file 1
$\{e, f\}$ denotes base station 2 chooses to cache file 2

$$c_a(f_a) = p_1 a_2 (1 - f_a)$$

$$c_b(f_b) = \begin{cases} p_2 a_2 & f_b = 2 \\ 0 & \text{else} \end{cases} \tag{45}$$

$$c_c(f_c) = p_3 a_2 (1 - f_c)$$

$$c_d(f_d) = a_1 p_1 (1 - f_d)$$

$$c_e(f_e) = \begin{cases} a_1 p_2 & f_e = 2 \\ 0 & \text{else} \end{cases} \tag{46}$$

$$c_f(f_f) = a_1 p_3 (1 - f_f)$$

B.1.3 3 players with capacity 2, 3 files

A base station can cache $\{1, 2\}, \{1, 3\}, \{2, 3\}$

$$s_1 = \{A^1_{1,2}, A^1_{2,3}, A^1_{1,3}\}$$
$$A^1_{1,2} = \{U^3_1, U^3_{1,2}, U^3_{1,3}, U^3_{1,2,3}\}$$
$$A^1_{2,3} = \{U^1_1, U^1_{1,2}, U^1_{1,3}, U^1_{1,2,3}\}$$
$$A^1_{1,3} = \{U^2_1, U^2_{1,2}, U^2_{1,3}, U^2_{1,2,3}\}$$

$$s_2 = \{A^2_{1,2}, A^2_{2,3}, A^2_{1,3}\}$$
$$A^2_{1,3} = \{U^2_2, U^2_{1,2}, U^2_{2,3}, U^2_{1,2,3}\}$$
$$A^2_{2,3} = \{U^1_2, U^1_{1,2}, U^1_{2,3}, U^1_{1,2,3}\}$$
$$A^2_{1,2} = \{U^3_2, U^3_{1,2}, U^3_{2,3}, U^3_{1,2,3}\}$$

$$s_3 = \{A^3_{1,2}, A^3_{2,3}, A^3_{1,3}\}$$
$$A^3_{1,2} = \{U^3_3, U^3_{1,3}, U^3_{2,3}, U^3_{1,2,3}\}$$

$$A^3_{1,3} = \{U^2_3, U^2_{1,3}, U^2_{2,3}, U^2_{1,2,3}\}$$
$$A^3_{2,3} = \{U^1_3, U^1_{1,3}, U^1_{2,3}, U^1_{1,2,3}\}$$

Assume $p_1 = A^1_{1,2}, p_2 = A^2_{2,3}, p_3 = A^3_{1,2}$, these three paths form the strategy profile.

$$c(f) = \sum_{i=1}^{3} \sum e \in p_i c_e(f_e)$$

$$\sum_{e\in p_1} c_e(f_e) = \sum_{e\in A^1_{1,2}} c_e(f_e) = c_{U^3_1} f(U^3_1) + c_{U^3_{1,2}} f(U^3_{1,2}) + c_{U^3_{1,2,3}}(f_{U^3_{1,2,3}})$$

$$c_{U^3_1}(f_{U^3_1}) = a_3 p_1 1_{\{f_{U^3_1}=1\}}$$
$$c_{U^3_{1,2}}(f_{U^3_{1,2}}) = a_3 p_{1,2} 1_{\{f_{U^3_{1,2}}=2\}}$$
$$c_{U^3_{1,3}}(f_{U^3_{1,3}}) = a_3 p_{1,3} 1_{\{f_{U^3_{1,3}}=2\}}$$
$$c_{U^3_{1,2,3}}(f_{U^3_{1,2,3}}) = a_3 p_{1,2,3} 1_{\{f_{U^3_{1,2,3}}=3\}}$$

B.1.4 Back to general case

Now we can obtain the cost function from what we've seen above and rewrite the equations.

$$c_e(f_e) = a_n p_T 1_{\{|f_e|\geq|T|\}}$$
$$c(f) = \sum_{e\in E} c_e(f_e) = \sum_{e\in E} f_e a_n p_T 1_{\{|f_e|\geq|T|\}}$$
$$c(f^*) = \sum_{e\in E} c_e(f^*_e) = \sum_{e\in E} f^*_e a_n p_T 1_{\{|f^*_e|\geq|T|\}}$$

Where $a_n = \frac{n^{-\gamma}}{\sum_{i=1}^n i^{-\gamma}}$, $e = U^n_T$, and we can immediately conclude that the cost function is non-decreasing.

$$\sum_{e\in E} f_e c_e(f_e) \leq \sum_{e\in E} f^*_e c_e(f_e + 1)$$
$$= \sum_{e\in E} f^*_e a_n p_T 1_{\{f_e\geq|T|-1\}}$$
$$= \sum_{e\in E,|T|=1} f^*_e a_n p_T + \sum_{e\in E,|T|>1} f^*_e a_n p_T 1_{\{f_e\geq|T|-1\}}$$

The last equation holds because of the following equation

$$f^*_e \leq |T| \leq 1$$

And the next inequality is also trivial

$$\sum_{e\in E,|T|=1} f^*_e a_n p_T \leq c(f^*)$$

Now we consider the $|T| \geq 1$ part

$$\sum_{e\in E,|T|>1} f^*_e a_n p_T 1_{\{f_e\geq|T|-1\}} \leq \sum_{e\in E,|T|>1} f^*_e \frac{f_e}{|T|-1} a_n p_T 1_{\{f_e\geq|T|-1\}}$$

Thus we get the entanglement of f_e, f^*_e, notice that

$$1_{\{f_e\geq|T|-1\}} = 1_{\{f_e=|T|-1\}} + 1_{\{f_e=|T|\}}$$

Apply $xy \leq \frac{1}{2}(x^2 + y^2)$

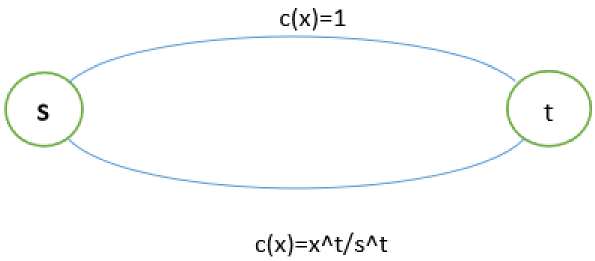


Figure 17: atomic pigou's game

$$f_e^* \frac{f_e}{|T|-1} a_n p_T 1_{\{f_e \geq |T|-1\}} \leq \sum_{e \in E, |T|>1} \frac{a_n p_T}{2(|T|-1)} (f_e^{*2} + f_e^2) 1_{\{f_e = |T|-1\}} + \sum_{e \in E} \frac{a_n p_T}{2(|T|-1)} (f_e^{*2} + f_e^2) 1_{\{f_e = |T|\}}$$
$$= \sum_{e \in E} [\frac{a_n p_T}{2(|T|-1)} f_e^{*2} 1_{\{f_e = |T|\}} + \frac{a_n p_T}{2(|T|-1)} f_e^2 1_{\{f_e = |T|\}} + \frac{a_n p_T}{2(|T|-1)} f_e^{*2} 1_{f_e = |T|-1} + \frac{a_n p_T}{2(|T|-1)} f_e^2 1_{f_e} \tag{47}$$

Where $\sum_{e\in E} \frac{a_n p_T}{2(|T|-1)} f_e^2 1_{\{f_e=|T|\}} = \sum_{e \in E} \frac{|T|}{2(|T|-1)} c_e(f)$, which has a beautiful form that we need, unfortunately with this method (at least within my ability), this is the "farthest" point we can reach. The most difficult part above is how to deal with the indicator function, we will consider applying m,n version (λ, μ) smoothness later.

B.1.5 Learn from "bad" example

"Always caching a resonable set of files" is the key as long as the capacity is not too small compared to the number of the files.

Atomic congestion game with arbitrary high PoA

For a non-atomic congestion game, we can find one with arbitrary high PoA easily, e.g. non-atomic pigou's game(we modified it to atomic version to obtain an example for comparing with our model in the following part). PoA in atomic selfish routing networks can be larger than in their nonatomic couterparts.

Important things need to be iterated for many times: In every atomic selfish routing network with affine cost functions, the POA is at most $\frac{5}{2}$. In atomic selfish routing networks with cost functions that are polynomial with nonnegative coefficients, the POA is at most a constant that depends on the maximum polynomial degree d , the dependence on d is exponential. For polynomial latency functions of degree d and pure strategies, we have $R = O(2^d d^{d+1})$,proof of this theorem includes loads of inequations.

Atomic pigou's game

Here is a simple derivative model of pigou's game with high PoA, we assume $s < N$(the number of players):

Atomic version Pigou's game with nonlinear variant. The higher the degree of the cost function is, the closer the capacity of the lower edge is to the number of the players and the higher the degree of the cost function, the higher the PoA might be. It can be easiliy observed that the worst-case equilibrium is: s players takes the lower edge and $N - s$ players take the upper edge, the total cost is N.

To obtain the optimal outcome, we suppose there are z players who take the lower edge, so the total cost is $cost = N - z + z\frac{z^t}{s^t}$, take derivative w.r.t z, and we can easily find that when $z = [\frac{s}{(t+1)^{\frac{1}{t}}}]$, the cost function reach its maximum.

$$cost_{opt} = N - \frac{s}{(t+1)^{\frac{1}{t}}} + \frac{[\frac{s}{(t+1)^{\frac{1}{t}}}]^{t+1}}{s^t}$$

Thus

$$PoA = \frac{N}{N - \frac{s}{(t+1)^{\frac{1}{t}}} + \frac{[\frac{s}{(t+1)^{\frac{1}{t}}}]^{t+1}}{s^t}}$$

$cost_{opt}$ $N = \frac{t}{t+1} \frac{s}{(t+1)^{\frac{1}{t}}}$, let $t \to \infty$,$PoA \to \infty$

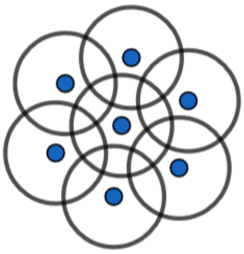


Figure 18: A base station has most 6 neighbors

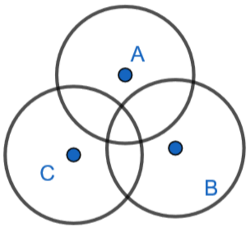


Figure 19: triangle model

An example with relatively high PoA construed by me

By sugar-water inequality, "independent" base station will decreass the PoA of the whole game (note the best startegy for an independent base station is always caching the most popular k files.), thus we can confine our discussion to intersecting base stations, for the same reason we consider the situations when two intersecting base stations have the largest common area.

Due to the assumption in the paper, the distance between any two base stations is larger than $\sqrt{2}$r, where r is the radius of the disc, thus a base station can intersect with at most six other base stations.

According to the paper, set Zipf parameter $\gamma = 1$, distance bertween caches is set to $d = r\sqrt{2}$ m, these setting s can give us an overall feeling about the relative value relationship, and help us to analyse the simple model.

We omit the calculation details here and get:

$$Area_1 = Area_2 = Area_3 = \frac{\pi}{4} + \frac{\sqrt{3}}{2} + 0.5$$

$$Area_{1,2} = Area_{1,3} = Area_{2,3} = \frac{\pi}{4} - \frac{\sqrt{3}}{2} + 0.5$$

$$Area_{1,2,3} = \frac{\sqrt{3}}{2} + \frac{\pi}{4} - 1.5$$

$$p_1 = p_2 = p_3 = 0.2736, p_{1,2} = p_{2,3} = p_{1,3} = 0.0533, p_{1,2,3} = 0.0193$$

$$a_1 = \frac{6}{11}, a_2 = \frac{3}{11}, a_3 = \frac{2}{11}$$

By enumeration method (note that base station 1,2,3 have the same status in our model), we obtain that the Nash equilibrium is

$$s_1^{NE} = \{1, 2\}, s_2^{NE} = \{1, 2\}, s_3^{NE} = \{1, 2\}$$

however the optimum strategy profile is

$$s_1^{opt} = \{1, 2\}, s_2^{opt} = \{1, 2\}, s_3^{opt} = \{1, 3\}$$

and $PoA \approx \frac{2.397}{2.1284} \approx 1.1262$.

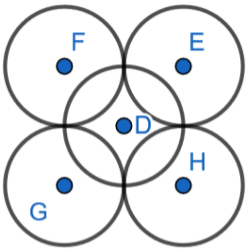


Figure 20: square model

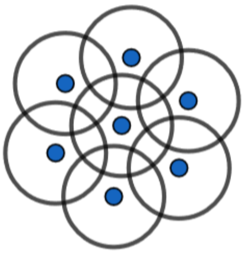


Figure 21: flower model

The above example has relatively higher PoA, when we increase the number of the files from 3 to 4, and keep the capacity to 2,we get $PoA = 1.0684$, when we increase the number of the files from 3 to 4, capacity from 2 to 3, we found that both the Nash equilibrium and optimal strategy profile is

$$s_1 = \{1, 2, 3\}, s_2 = \{1, 2, 3\}, s_3 = \{1, 2, 4\}$$

so the PoA=1 in this situation.

In the square model, four files, capacity is two, the optimum and the Nash equilibirum outcome are the same:

$$s_D = \{2, 4\}, s_E = \{1, 2\}, s_F = \{1, 2\}, s_G = \{1, 2\}, s_H = \{1, 2\}$$

To modify γ is not our first choice because it depends on the exsitent phenomenon in the real world.

Analysis: differences

In pigou's game, to obtain the optimal outcome, there has to be a few selfless players, these abject players need to endure the relatively higher cost when most of other players are enjoying the fantastic shortcut. And their insistence is crucial to the whole community because as soon as they join the party of taking the shortcut, the shortcut is no more a shortcut, the situation is abysmal for everyone all over again.

However, in our model, if a base station hope to serve as many as possible other players (in our case the maximal number is 6), it has to intersect as much as possible with other base stations, image a situation similar to the "pigou's tragedy" above, there have to be a selfless base station who has much larger miss probability in the region which is covered by itself than in its intersecting region.

In other words, in our model: You won't be too miserable when your neighbors live happy lives. You get some returns as you choose to be selfless and more people , and when you don't have many neighbors, independence makes "selfishness" more in correspondence with "optimal". Now we have a second look at the six-petal flower model mentioned before, the intersecting area(including 2-intersecting and 3-intersecting) account for approximately 80% area of the whole disc.

The main task is to quantify the "happy", "miserable","selfless". What a selfless base station can do is decreasing its neighbors remarkably, consider the different intersecting area ratios. Now a model of three base stations and the upper bound of PoA is demanded.

B.1.6 m,n version smoothness inequality doesn't work

Brief description on the method

Generally, we define the $(\lambda - \mu)$ smoothness by a summation form inequality:

$$\sum_{i=1}^{k} c_i(s_i^*, s_{-i}) \leq \lambda cost(s^*) + \mu cost(s)$$

for every pair s, s^*.

$(\lambda - \mu)$ smoothness has a sufficient condition: $\forall 0 \leq n \leq N, \forall 0 \leq m \leq N$

$$\lambda m c_e(m) \geq m c_e(n+1) - \mu n c_e(n)$$

The proof of this sufficient condition includes two keys:

1. sum over $e \in E$.
2. $\sum_{e \in E} f_e(s^*) c_e(f_e(s)+1) \geq \sum_{e \in E} f_e(s) c_e(f_e(s))$ where s is the NE outcome.

Why it doesn't work

However, as we can see from the expression of $c_e(f_e)$ in our model

$$c_e(f_e) = a_n p_T 1_{\{|f_e| \geq |T|\}}$$

where a_n is the probability of the user requests file n and p_T is the probability of the user appears within region T, f_e is the number of players on the resource e.

After observation one can find that there is a jump of the value from 0 to 1 in this function. Let $n = N-1, m = N-2$ thus $LHS = 0$ while $RHS = (N-2)a_n p_T$ so there can't be a pair of (λ, μ) to make this inequation hold for all $m, n \leq N$(the number of the players), which means when $N > 2$ this inequation does not hold for any λ, μ, thus we need to try another way or another equivalent game.

$$\lambda m c_e(m) \geq m c_e(n+1) - \mu n c_e(n), \forall 0 \leq n \leq N, \forall 0 \leq m \leq N$$

B.2 Market Sharing Game Model

In market sharing game, agents provide service to markets, agents are the players, each agent j has a total budget B_j, each agent should decide which subset of markets to service. Agent j can service a subset S_j of markets if the total cost is within its budget.

Each market has a query rate q_i, the rate at which market i is requested per unit time, each market i also has a cost C_i corresponding to the cost for serving this market

Counterparts: Market Sharing Game V.S. Congestion Game

You can find more about utility system and market sharing game in [3].

- Agents or players correspond to base stations
- Markets correspond to files
- Query rate of market i is the query rate of file i: a_i.
- The cost C_i corresponds to the size of file i
- Total budget B_j of agent j corresponds to the capacity
- The markets of the market sharing game correspond to the resources

Here are the settings: There are n players(i=1,...,n), t primary factors(k=1,...,t), the i^{th} player's(i=1,...,n) set of pure strategies contain s_i elements($r_i = 1, ..., s_i$). The r_i^{th} pure strategy may be viewed as the selection of a particular subset of the costs of each of the resources he selects.

Market Sharing Game is a valid Utility System. The conditions for a valid utility system are :

- The social utility and private utility functions are measured in the same standard unit
- The social utility function is submodular

- The private utility of an agent is at least the change in social utility that would occur if the agent declined to participate in the game
- The sum of the private utilities of the agents is at most the social utility.
- A valid utility system has an upper bound 2 for PoA and many good characteristics.

Failed to modify $f^{(m)}(B)$ and failed (decline to participate in the game part goes wrong) and the relaxed conclusion is not applicable for our model.

B.3 Wrong way to construe derivative inequality

Construing λ, μ to make the inequality hold:

$$\lambda c(s^*) \geq \sum_{i=1}^{N}(c_i(s_i^*, s_{-i}) - c_i(s)) + (1-\mu)c(s)$$

On the LHS we have $\lambda c(s^*)$ where we bound "$c(s^*)$" with "$\sum_{i=1}^{N} c_i(s^*)$". On the RHS we bound "$c(s)$" with "$\sum_{i=1}^{N} c_i(s)$". However if we use the inequality , and the items from the LHS and RHS don't correspond, so I apply some other inequalities.

First, let $1-\mu = 1$, thus $-\sum_{i=1}^{2} c_i(s) + c(s) \leq 0$.

$$c(s^*) \geq \frac{2(a_{k+1} + ... + a_{2k})(1-A)}{2-A}$$

This inequality assume the best situation when all file can be accessed in the intersecting region and the other regions only miss the least popular files.

$$\begin{aligned}\sum_{i=1}^{2} c_i(s_i^*, s_{-i}) &\leq \sum_{l\in\{1,2,...,2k\}\backslash\{i_1^*,i_2^*,...,i_k^*\}} \frac{a_l}{2-A} + \sum_{q\in\{1,2,...,2k\}\backslash\{j_1^*,j_2^*,...,j_k^*\}} \frac{a_q}{2-A} \\ &\leq \sum_{s\in\{1,2,...,2k\}\backslash\{i_1^*,i_2^*,...,i_k^*\}\cup\{1,2,...,2k\}\backslash\{j_1^*,j_2^*,...,j_k^*\}} \frac{a_s}{2-A} + \sum_{t\in\{1,2,...,2k\}\backslash\{i_1^*,i_2^*,...,i_k^*\}\cap\{1,2,...,2k\}\backslash\{j_1^*,j_2^*,...,j_k^*\}} \frac{a_t}{2-A} \\ &\leq \frac{1 + a_{k+1} + ... + a_{2k}}{2-A}\end{aligned} \tag{48}$$

This inequality uses the fact that for optimal strategies and Nash equilibriums $\{1, 2, ..., k\} \subset \{i_1, i_2, ..., i_k\} \cup \{j_1, j_2, ..., j_k\}$, where numbers represent files.

Let $\lambda = \frac{1+a_{k+1}+...+a_{2k}}{2(a_{k+1}+...+a_{2k})(1-A)}$, thus $\frac{\lambda}{1-\mu} = \frac{1+a_{k+1}+...+a_{2k}}{2(a_{k+1}+...+a_{2k})(1-A)}$.