Hedonic Clustering Games

Moran Feldman, Liane Lewin-Eytan and Joseph (Seffi) Naor

Nome: Felipe L. Galvão
RA:116790

June 22, 2017

Introdução

Introdução

O que é um jogo hedônico?

Introdução

O que é um jogo hedônico?

Introdução

O que é um jogo hedônico?

- Conceito introduzido por Drèze e Greenberg [DG80] para jogos de formação de coalizão

Introdução

O que é um jogo hedônico?

- Conceito introduzido por Drèze e Greenberg [DG80] para jogos de formação de coalizão
- Jogos em que a utilidade do jogador não depende da organização de outras coalizões

Introdução

O que é um jogo hedônico?

- Conceito introduzido por Drèze e Greenberg [DG80] para jogos de formação de coalizão
- Jogos em que a utilidade do jogador não depende da organização de outras coalizões
- A literatura existente aborda jogos hedônicos cooperativos e não-cooperativos, existência de coalizões estáveis (núcleo ou individual), complexidade para encontrar tais equilíbrios

Introdução

O que é um jogo hedônico?

- Conceito introduzido por Drèze e Greenberg [DG80] para jogos de formação de coalizão
- Jogos em que a utilidade do jogador não depende da organização de outras coalizões
- A literatura existente aborda jogos hedônicos cooperativos e não-cooperativos, existência de coalizões estáveis (núcleo ou individual), complexidade para encontrar tais equilíbrios

O que é clustering?

Introdução

O que é um jogo hedônico?

- Conceito introduzido por Drèze e Greenberg [DG80] para jogos de formação de coalizão
- Jogos em que a utilidade do jogador não depende da organização de outras coalizões
- A literatura existente aborda jogos hedônicos cooperativos e não-cooperativos, existência de coalizões estáveis (núcleo ou individual), complexidade para encontrar tais equilíbrios

O que é clustering?

- Agrupamento de elementos por uma medida de similaridade (e.g. distância)

Introdução

O que é um jogo hedônico?

- Conceito introduzido por Drèze e Greenberg [DG80] para jogos de formação de coalizão
- Jogos em que a utilidade do jogador não depende da organização de outras coalizões
- A literatura existente aborda jogos hedônicos cooperativos e não-cooperativos, existência de coalizões estáveis (núcleo ou individual), complexidade para encontrar tais equilíbrios

O que é clustering?

- Agrupamento de elementos por uma medida de similaridade (e.g. distância)
- Tradicionalmente tratado como um problema de otimização

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot, \cdot)$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot, \cdot)$
- k clusters

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot, \cdot)$
- k clusters

Cada jogador:

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot,\cdot)$
- k clusters

Cada jogador:

- Controla um ponto

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot, \cdot)$
- k clusters

Cada jogador:

- Controla um ponto
- Escolhe em qual cluster aloca o ponto

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot, \cdot)$
- k clusters

Cada jogador:

- Controla um ponto
- Escolhe em qual cluster aloca o ponto
- Conjunto de estratégia do jogador i é $S_i = [k]$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Dados:

- n pontos num espaço métrico com função de distância $d(\cdot, \cdot)$
- k clusters

Cada jogador:

- Controla um ponto
- Escolhe em qual cluster aloca o ponto
- Conjunto de estratégia do jogador i é $S_i = [k]$

Iremos nos referir aos jogadores diretamente pelos pontos $u, v, w \ldots \in [n]$.

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

As escolhas dos jogadores geram uma atribuição de pontos aos clusters:

$$A : [n] \to [k]$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

As escolhas dos jogadores geram uma atribuição de pontos aos clusters:

$$A : [n] \to [k]$$

Cada cluster constitui um conjunto (coalizão):

$$C_j = \{u \in [n] \mid A(u) = j\} \quad j = 1 \dots k$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

As escolhas dos jogadores geram uma atribuição de pontos aos clusters:

$$A : [n] \to [k]$$

Cada cluster constitui um conjunto (coalizão):

$$C_j = \{u \in [n] \mid A(u) = j\} \quad j = 1 \dots k$$

Definimos a função $D(u, C)$ como a distância de um ponto a todos os elementos de um cluster:

$$D(u, C) = \sum_{v \in C} d(u, v)$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Cada cluster C tem um centroide dado pela função $h : [k] \to [n]$:

$$h(C) = \arg\min_{u \in C} \{D(u, C)\}$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Cada cluster C tem um centroide dado pela função $h : [k] \to [n]$:

$$h(C) = \arg\min_{u \in C} \{D(u, C)\}$$

O custo do jogador u é dado pela sua distância ao centroide de seu cluster:

$$c(u) = d(u, h(A(u)))$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Cada cluster C tem um centroide dado pela função $h : [k] \rightarrow [n]$:

$$h(C) = \arg\min_{u \in C} \{D(u, C)\}$$

O custo do jogador u é dado pela sua distância ao centroide de seu cluster:

$$c(u) = d(u, h(A(u)))$$

Do qual derivamos o custo de um cluster:

$$c(C) = \sum_{u \in C} c(u) = D(h(C), C) = D(C)$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

Cada cluster C tem um centroide dado pela função $h : [k] \to [n]$:

$$h(C) = \arg\min_{u \in C} \{D(u, C)\}$$

O custo do jogador u é dado pela sua distância ao centroide de seu cluster:

$$c(u) = d(u, h(A(u)))$$

Do qual derivamos o custo de um cluster:

$$c(C) = \sum_{u \in C} c(u) = D(h(C), C) = D(C)$$

E o custo social:

$$CS = \sum_{u \in [n]} c(u) = \sum_{j=1}^{k} c(C_j)$$

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

- Jogadores egoístas querendo escolher um cluster de forma a minimizar seu custo.

Jogo de clustering hedônico

Clustering fixo - modelo das k-medianas

- Jogadores egoístas querendo escolher um cluster de forma a minimizar seu custo.
- Estamos interessados em encontrar equilíbrios.

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Dado que C^{+v} representa um cluster com a adição do jogador v, um resultado de clustering $C_1, C_2, \ldots, C_k$ é um equilíbrio puro de Nash se:

$$d(u, h(C_j^{+v})) \geq d(u, h(A(u))) \quad \forall u \in [n], j = 0 \ldots k$$

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Dado que C^{+v} representa um cluster com a adição do jogador v, um resultado de clustering $C_1, C_2, \ldots, C_k$ é um equilíbrio puro de Nash se:

$$d(u, h(C_j^{+v})) \geq d(u, h(A(u))) \quad \forall u \in [n], j = 0 \ldots k$$

Ou seja, nenhum jogador consegue reduzir seu custo movendo para outro cluster.

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Dado que C^{+v} representa um cluster com a adição do jogador v, um resultado de clustering $C_1, C_2, \ldots, C_k$ é um equilíbrio puro de Nash se:

$$d(u, h(C_j^{+v})) \geq d(u, h(A(u))) \quad \forall u \in [n], j = 0 \ldots k$$

Ou seja, nenhum jogador consegue reduzir seu custo movendo para outro cluster.

Tal como definimos h, o centroide não é único. Precisamos de um critério de desempate.

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Num jogo puramente hedônico o critério de desempate precisa ser estático dado os membros que constituem cada cluster.

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Num jogo puramente hedônico o critério de desempate precisa ser estático dado os membros que constituem cada cluster.

Proposição 1

Com critério de desempate estático, um equilíbrio de Nash pode não existir.

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Prova

Para $[n] = \{u, v, w\}$, $k = 2$ e o seguinte critério de desempate estático:

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Prova

Para $[n] = \{u, v, w\}$, $k = 2$ e o seguinte critério de desempate estático:

Pontos do cluster	Posição do centroide
$\{u, v\}$	u
$\{v, w\}$	v
$\{w, u\}$	w

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Prova

Para $[n] = \{u, v, w\}$, $k = 2$ e o seguinte critério de desempate estático:

Pontos do cluster	Posição do centroide
$\{u, v\}$	u
$\{v, w\}$	v
$\{w, u\}$	w

Não existe equilíbrio independente das posições u, v e w. $\square$

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Para evitar esse problema introduzimos um critério de desempate dependente do histórico:

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Para evitar esse problema introduzimos um critério de desempate dependente do histórico:

- Inicializamos o jogo com um critério de desempate estático R qualquer

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Para evitar esse problema introduzimos um critério de desempate dependente do histórico:

- Inicializamos o jogo com um critério de desempate estático R qualquer
- Dado que um jogador muda de estratégia, os centroides continuam sendo os mesmos se seus pontos ainda forem posições válidas para o centroide (i.e. minimizam o custo do cluster)

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Para evitar esse problema introduzimos um critério de desempate dependente do histórico:

- Inicializamos o jogo com um critério de desempate estático R qualquer
- Dado que um jogador muda de estratégia, os centroides continuam sendo os mesmos se seus pontos ainda forem posições válidas para o centroide (i.e. minimizam o custo do cluster)
- Caso contrário, novos centroides são definidos baseados em R

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Para evitar esse problema introduzimos um critério de desempate dependente do histórico:

- Inicializamos o jogo com um critério de desempate estático R qualquer
- Dado que um jogador muda de estratégia, os centroides continuam sendo os mesmos se seus pontos ainda forem posições válidas para o centroide (i.e. minimizam o custo do cluster)
- Caso contrário, novos centroides são definidos baseados em R

Pequeno desvio de jogos hedônicos mas ainda podemos tirar conclusões sobre critérios de desempate estáticos baseados no critério de desempate dependente do histórico.

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Vamos mostrar que:

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Vamos mostrar que:

- O equilíbrio pode não existir no caso geral

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Vamos mostrar que:

- O equilíbrio pode não existir no caso geral
- Como forçar um equilíbrio impondo uma penalidade para mudanças de estratégia que alterem os centroides, e uma forma natural de definir tal penalidade

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Vamos mostrar que:

- O equilíbrio pode não existir no caso geral
- Como forçar um equilíbrio impondo uma penalidade para mudanças de estratégia que alterem os centroides, e uma forma natural de definir tal penalidade
- Existe equilíbrio para os casos de espaço métrico em linha e em árvore, mesmo sem penalidade

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Vamos mostrar que:

- O equilíbrio pode não existir no caso geral
- Como forçar um equilíbrio impondo uma penalidade para mudanças de estratégia que alterem os centroides, e uma forma natural de definir tal penalidade
- Existe equilíbrio para os casos de espaço métrico em linha e em árvore, mesmo sem penalidade
- Para o espaço métrico em linha, jogadores jogando respostas ótimas sempre convergem para um equilíbrio

Equilíbrio puro de Nash

Clustering fixo - modelo das k-medianas

Vamos mostrar que:

- O equilíbrio pode não existir no caso geral
- Como forçar um equilíbrio impondo uma penalidade para mudanças de estratégia que alterem os centroides, e uma forma natural de definir tal penalidade
- Existe equilíbrio para os casos de espaço métrico em linha e em árvore, mesmo sem penalidade
- Para o espaço métrico em linha, jogadores jogando respostas ótimas sempre convergem para um equilíbrio
- Em todos os três casos de equilíbrio mencionados, o preço da estabilidade é 1 e o preço da anarquia é ilimitado

Caso sem equilíbrio

Clustering fixo - modelo das k-medianas - caso geral

Teorema 1

Existe uma instância do problema com 9 pontos e 2 clusters sem nenhum equilíbrio puro de Nash

Caso sem equilíbrio

Clustering fixo - modelo das k-medianas - caso geral

Teorema 1

Existe uma instância do problema com 9 pontos e 2 clusters sem nenhum equilíbrio puro de Nash

A prova é extensa, fazendo uma análise por casos de uma instância específica para mostrar que não há equilíbrio. Mostraremos apenas a intuição da prova.

Caso sem equilíbrio

Clustering fixo - modelo das k-medianas - caso geral

Representamos a instância por um grafo onde o peso das arestas definem distâncias:

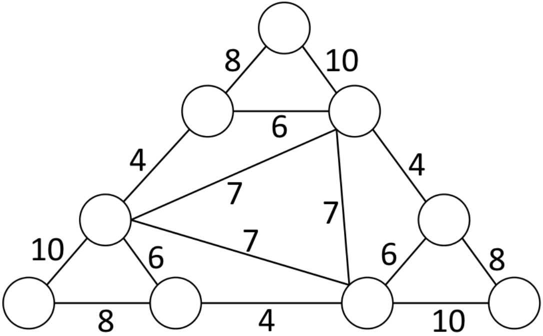

Figure: Instância sem equilíbrio

Caso sem equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Analisamos cada possível par de centroides, associando cada nó ao centroide mais próximo e verificando se:

- existe mudança de centroide que reduz o custo do cluster
- algum jogador melhora seu custo mudando de cluster

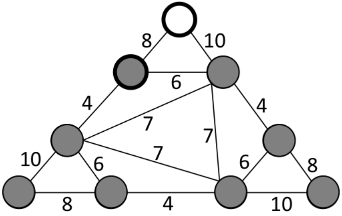

Figure: Exemplo de par de centroides

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Intuitivamente, se penalizarmos o suficiente mudanças de estratégia que alterem os centroides iremos forçar um equilíbrio. Como determinar essa penalidade?

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Intuitivamente, se penalizarmos o suficiente mudanças de estratégia que alterem os centroides iremos forçar um equilíbrio. Como determinar essa penalidade?

$\rightarrow$ Penalidade baseada na externalidade que causa (similar a um mecanismo VCG).

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Além do seu custo convencional, se um jogador v deseja mover para um cluster C ele paga um valor igual à distância entre o centroide de C antes e depois de acrescentar v:

$$c(v) = d(v, h(A(v))) + d(h(C), h(C^{+v}))$$

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Além do seu custo convencional, se um jogador v deseja mover para um cluster C ele paga um valor igual à distância entre o centroide de C antes e depois de acrescentar v:

$$c(v) = d(v, h(A(v))) + d(h(C), h(C^{+v}))$$

Lema 1

Um jogador u só irá mudar sua estratégia de um cluster C_1 para C_2 se sua distância do centroide original $h(C_2)$ for menor que sua distância de $h(C_1)$.

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Prova

Mudar para C_2 é resposta ótima quando:

$$d(u, h(C_2^{+v})) + d(h(C_2^{+v}), h(C_2)) < d(u, h(C_1))$$

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Prova

Mudar para C_2 é resposta ótima quando:

$$d(u, h(C_2^{+v})) + d(h(C_2^{+v}), h(C_2)) < d(u, h(C_1))$$

E por desigualdade triangular sabemos que:

$$d(u, h(C_2)) \leq d(u, h(C_2^{+v})) + d(h(C_2^{+v}), h(C_2))$$

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Prova

Mudar para C_2 é resposta ótima quando:

$$d(u, h(C_2^{+v})) + d(h(C_2^{+v}), h(C_2)) < d(u, h(C_1))$$

E por desigualdade triangular sabemos que:

$$d(u, h(C_2)) \leq d(u, h(C_2^{+v})) + d(h(C_2^{+v}), h(C_2))$$

Juntando ambas as desigualdades concluímos:

$$d(u, h(C_2)) < d(u, h(C_1))$$

$\square$

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Teorema 2

Qualquer resposta ótima estritamente reduz o custo social quando aplicamos essa penalidade.

Forçando um equilíbrio

Clustering fixo - modelo das *k*-medianas - caso geral com penalidade

Teorema 2

Qualquer resposta ótima estritamente reduz o custo social quando aplicamos essa penalidade.

Como o jogo é finito, isso implica que sempre existe equilíbrio de Nash e que sempre chegamos nesse equilíbrio se os jogadores apenas jogarem respostas ótimas.

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Prova

Analisamos o resultado de uma resposta ótima de um jogador u em dois passos:

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Prova

Analisamos o resultado de uma resposta ótima de um jogador u em dois passos:

1. Sua mudança de um cluster C_1 para C_2. Pelo lema 1 e considerando que nenhum outro jogador é afetado, o custo social necessariamente diminui

Forçando um equilíbrio

Clustering fixo - modelo das *k*-medianas - caso geral com penalidade

Prova

Analisamos o resultado de uma resposta ótima de um jogador u em dois passos:

1. Sua mudança de um cluster C_1 para C_2. Pelo lema 1 e considerando que nenhum outro jogador é afetado, o custo social necessariamente diminui
2. A atualização dos centroides. Por definição o centroide é escolhido de forma a minimizar o custo do cluster, então uma mudança de centroide não pode aumentar o custo.

Forçando um equilíbrio

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Prova

Analisamos o resultado de uma resposta ótima de um jogador u em dois passos:

1. Sua mudança de um cluster C_1 para C_2. Pelo lema 1 e considerando que nenhum outro jogador é afetado, o custo social necessariamente diminui
2. A atualização dos centroides. Por definição o centroide é escolhido de forma a minimizar o custo do cluster, então uma mudança de centroide não pode aumentar o custo.

Como o passo 1 estritamente melhora o custo social e o passo 2 não piora, a resposta ótima estritamente melhora o custo social. □

Preço da estabilidade

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Corolário 1

O preço da estabilidade com essa penalidade é 1.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - caso geral com penalidade

Corolário 1

O preço da estabilidade com essa penalidade é 1.

Prova

Pelo teorema 2, qualquer resposta ótima diminui o custo social. Logo, qualquer solução ótima também precisa ser equilíbrio de Nash.

Preço da anarquia

Clustering fixo - modelo das k-medianas

Proposição 2

O preço da anarquia é ilimitado em todos os espaços métricos tratados.

Preço da anarquia

Clustering fixo - modelo das k-medianas

Proposição 2

O preço da anarquia é ilimitado em todos os espaços métricos tratados.

Prova

Considere o jogo numa linha com os pontos $[n] = \{(0), (1), (M)\}$, M arbitrariamente grande, e com $k = 2$. Temos os seguintes equilíbrios:

Preço da anarquia

Clustering fixo - modelo das k-medianas

Proposição 2

O preço da anarquia é ilimitado em todos os espaços métricos tratados.

Prova

Considere o jogo numa linha com os pontos $[n] = \{(0), (1), (M)\}$, M arbitrariamente grande, e com $k = 2$. Temos os seguintes equilíbrios:

- Clusters $\{(0)\}$ e $\{(1), (M)\}$, centroide do segundo cluster em (1). Pior equilíbrio de custo $M - 1$.

Preço da anarquia

Clustering fixo - modelo das k-medianas

Proposição 2

O preço da anarquia é ilimitado em todos os espaços métricos tratados.

Prova

Considere o jogo numa linha com os pontos $[n] = \{(0), (1), (M)\}$, M arbitrariamente grande, e com $k = 2$. Temos os seguintes equilíbrios:

- Clusters $\{(0)\}$ e $\{(1), (M)\}$, centroide do segundo cluster em (1). Pior equilíbrio de custo $M - 1$.
- Clusters $\{(M)\}$ e $\{(0), (1)\}$. Equilíbrio ótimo de custo 1.

Preço da anarquia

Clustering fixo - modelo das k-medianas

Proposição 2

O preço da anarquia é ilimitado em todos os espaços métricos tratados.

Prova

Considere o jogo numa linha com os pontos $[n] = \{(0), (1), (M)\}$, M arbitrariamente grande, e com $k = 2$. Temos os seguintes equilíbrios:

- Clusters $\{(0)\}$ e $\{(1), (M)\}$, centroide do segundo cluster em (1). Pior equilíbrio de custo $M - 1$.
- Clusters $\{(M)\}$ e $\{(0), (1)\}$. Equilíbrio ótimo de custo 1.

Preço da anarquia é $M - 1$. $\square$

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

No espaço métrico em linha, o centroide assume o valor convencional de mediana (valor central).

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

No espaço métrico em linha, o centroide assume o valor convencional de mediana (valor central).

Lema 2

Num espaço métrico em linha, qualquer resposta ótima estritamente diminui o custo social.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

No espaço métrico em linha, o centroide assume o valor convencional de mediana (valor central).

Lema 2

Num espaço métrico em linha, qualquer resposta ótima estritamente diminui o custo social.

Assumindo que um jogador v move de um cluster C_1 para um cluster C_2 para melhorar seu custo, escrevemos:

- C_1^{-v} para indicar o cluster C_1 após a saída de v
- C_2^{+v} para indicar o cluster C_2 após v se juntar

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

Prova

Iremos mostrar que os nós afetados pela mudança de estratégia do jogador v não pioram o custo social

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

Prova

Iremos mostrar que os nós afetados pela mudança de estratégia do jogador v não pioram o custo social

- Nós em C_1: Por definição do centroide, $c(C_1^{-v}) \leq c(C_1)$

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

Prova

Iremos mostrar que os nós afetados pela mudança de estratégia do jogador v não pioram o custo social

- Nós em C_1: Por definição do centroide, $c(C_1^{-v}) \leq c(C_1)$
- Nó v: A mudança de estratégia implica que v estritamente diminui seu custo, i.e. $d(v, h(C_2^{+v})) < d(v, h(C_1))$

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

- Nós em C_2: Temos dois casos

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

- Nós em C_2: Temos dois casos
  - Se o número de nós em C_2 é ímpar: C_2^{+v} será par e com duas medianas incluindo $h(C_2)$ que pelo nosso critério de desempate implica que $h(C_2^{+v}) = h(C_2)$

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

- Nós em C_2: Temos dois casos
  - Se o número de nós em C_2 é ímpar: C_2^{+v} será par e com duas medianas incluindo $h(C_2)$ que pelo nosso critério de desempate implica que $h(C_2^{+v}) = h(C_2)$
  - Se o número de nós em C_2 é par temos duas potenciais medianas que chamaremos de $h_L(C_2)$ e $h_R(C_2)$ para a da esquerda e direita, respectivamente.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

- Nós em C_2: Temos dois casos
  - Se o número de nós em C_2 é ímpar: C_2^{+v} será par e com duas medianas incluindo $h(C_2)$ que pelo nosso critério de desempate implica que $h(C_2^{+v}) = h(C_2)$
  - Se o número de nós em C_2 é par temos duas potenciais medianas que chamaremos de $h_L(C_2)$ e $h_R(C_2)$ para a da esquerda e direita, respectivamente.
    Observamos que $D(h_L(C_2), C_2) = D(h_R(C_2), C_2)$.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

- Nós em C_2: Temos dois casos
  - Se o número de nós em C_2 é ímpar: C_2^{+v} será par e com duas medianas incluindo $h(C_2)$ que pelo nosso critério de desempate implica que $h(C_2^{+v}) = h(C_2)$
  - Se o número de nós em C_2 é par temos duas potenciais medianas que chamaremos de $h_L(C_2)$ e $h_R(C_2)$ para a da esquerda e direita, respectivamente.
    Observamos que $D(h_L(C_2), C_2) = D(h_R(C_2), C_2)$.
    Se v ficar a esquerda de $h_L(C_2)$ ou a direita de $h_R(C_2)$ continuará sendo uma das duas anteriores.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

- Nós em C_2: Temos dois casos
  - Se o número de nós em C_2 é ímpar: C_2^{+v} será par e com duas medianas incluindo $h(C_2)$ que pelo nosso critério de desempate implica que $h(C_2^{+v}) = h(C_2)$
  - Se o número de nós em C_2 é par temos duas potenciais medianas que chamaremos de $h_L(C_2)$ e $h_R(C_2)$ para a da esquerda e direita, respectivamente.
    Observamos que $D(h_L(C_2), C_2) = D(h_R(C_2), C_2)$.
    Se v ficar a esquerda de $h_L(C_2)$ ou a direita de $h_R(C_2)$ continuará sendo uma das duas anteriores.
    Resta então analisar o caso de v estar entre $h_L(C_2)$ e $h_R(C_2)$, se tornando a nova mediana.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

Nesse caso temos que:

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

Nesse caso temos que:

$$D(v, C_2) = \sum_{u \in C_2} d(v, u)$$

$$= \sum_{u \in C_2} [\frac{d(h_L(C_2), v)}{d(h_L(C_2), h_R(C_2))} \cdot d(v, h_L(C_2)) + \frac{d(v, h_L(C_1))}{d(h_L(C_2), h_R(C_2))} \cdot d(v, h_R(C_2))$$

$$= \frac{d(h_L(C_2), v)}{d(h_L(C_2), h_R(C_2))} \cdot D(h_L(C_2), C_2) + \frac{d(v, h_L(C_1))}{d(h_L(C_2), h_R(C_2))} \cdot D(h_R(C_2, C_2))$$

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em linha

Nesse caso temos que:

$$D(v, C_2) = \sum_{u \in C_2} d(v, u)$$

$$= \sum_{u \in C_2} [\frac{d(h_L(C_2), v)}{d(h_L(C_2), h_R(C_2))} \cdot d(v, h_L(C_2)) + \frac{d(v, h_L(C_1))}{d(h_L(C_2), h_R(C_2))} \cdot d(v, h_R(C_2))$$

$$= \frac{d(h_L(C_2), v)}{d(h_L(C_2), h_R(C_2))} \cdot D(h_L(C_2), C_2) + \frac{d(v, h_L(C_1))}{d(h_L(C_2), h_R(C_2))} \cdot D(h_R(C_2, C_2))$$

E, como $D(h_L(C_2), C_2) = D(h_R(C_2), C_2)$, o último valor da equação de iguala a eles.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

Resultados dependem de demonstrar algumas propriedades específicas desse tipo de espaço métrico. Iremos apenas enunciá-los.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

Definição 1

Dado uma árvore T com $|V|$ nós, um nó $v^\star \in T$ é chamado de nó mediano se sua remoção gera componentes conexos com tamanho máximo $\lceil \frac{|V|-1}{2} \rceil$ cada.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

Definição 1

Dado uma árvore T com $|V|$ nós, um nó $v^\star \in T$ é chamado de nó mediano se sua remoção gera componentes conexos com tamanho máximo $\lceil \frac{|V|-1}{2} \rceil$ cada.

Lema 3

Há no máximo 2 nós medianos numa árvore.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

Definição 1

Dado uma árvore T com $|V|$ nós, um nó $v^\star \in T$ é chamado de nó mediano se sua remoção gera componentes conexos com tamanho máximo $\lceil \frac{|V|-1}{2} \rceil$ cada.

Lema 3

Há no máximo 2 nós medianos numa árvore.

Definição 2

Um cluster C tem a propriedade de fechamento caso todo os nós num caminho entre elementos de C também façam parte de C.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

Definição 1

Dado uma árvore T com $|V|$ nós, um nó $v^\star \in T$ é chamado de nó mediano se sua remoção gera componentes conexos com tamanho máximo $\lceil \frac{|V|-1}{2} \rceil$ cada.

Lema 3

Há no máximo 2 nós medianos numa árvore.

Definição 2

Um cluster C tem a propriedade de fechamento caso todo os nós num caminho entre elementos de C também façam parte de C.

Lema 4

O centroide de um cluster com a propriedade de fechamento é sempre um nó mediano.

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

É possível demonstrar que existe um conjunto específico de respostas ótimas que:

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

É possível demonstrar que existe um conjunto específico de respostas ótimas que:

- Mantém a propriedade de fechamento (permitindo a caracterização dos centroides como medianas)

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

É possível demonstrar que existe um conjunto específico de respostas ótimas que:

- Mantém a propriedade de fechamento (permitindo a caracterização dos centroides como medianas)
- Estritamente diminuem o custo social

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

É possível demonstrar que existe um conjunto específico de respostas ótimas que:

- Mantém a propriedade de fechamento (permitindo a caracterização dos centroides como medianas)
- Estritamente diminuem o custo social

Caso a configuração inicial não gere clusters com a propriedade de fechamento, é sempre possível chegar em uma. Isso nos dá o resultado mais geral:

Preço da estabilidade

Clustering fixo - modelo das k-medianas - espaço métrico em árvore

É possível demonstrar que existe um conjunto específico de respostas ótimas que:

- Mantém a propriedade de fechamento (permitindo a caracterização dos centroides como medianas)
- Estritamente diminuem o custo social

Caso a configuração inicial não gere clusters com a propriedade de fechamento, é sempre possível chegar em uma. Isso nos dá o resultado mais geral:

Teorema 3

Sempre existe equilíbrio de Nash no caso de espaços métricos em árvore.

Corolário

O preço da estabilidade nesse caso é 1.

Conclusão - Outros modelos

Clustering fixo - modelo dos k-centros

Mudamos a função que seleciona o centroide para escolher o elemento que minimiza o raio do cluster:

Conclusão - Outros modelos

Clustering fixo - modelo dos k-centros

Mudamos a função que seleciona o centroide para escolher o elemento que minimiza o raio do cluster:

$$h(C) = \arg\min_{u \in C}\{\max_{v \in C} d(u, , v)\}$$

Conclusão - Outros modelos

Clustering de correlação

Um jogador tem a opção de criar novos clusters.

Conclusão - Outros modelos

Clustering de correlação

Um jogador tem a opção de criar novos clusters.

A medida de distância passa a ser diretamente uma medida de similaridade entre elementos $d(\cdot, \cdot) \in [0, 1]$.

Conclusão - Outros modelos

Clustering de correlação

Um jogador tem a opção de criar novos clusters.
A medida de distância passa a ser diretamente uma medida de similaridade entre elementos $d(\cdot, \cdot) \in [0, 1]$.
Cada ponto v tem um peso w_v que mede sua influência sobre outros elementos.

Conclusão

Usos no mundo real

Clustering fixo é usado para modelar uma rede de sensores formada por um grande número de aparelhos autônomos que investem recursos para comunicarem entre si.

Conclusão

Usos no mundo real

Clustering fixo é usado para modelar uma rede de sensores formada por um grande número de aparelhos autônomos que investem recursos para comunicarem entre si.
Clustering de correlação é usado para modelar editoras que querem se juntar para pagar por um serviço de marketing.

References I

- Jacques H Dreze and Joseph Greenberg, *Hedonic coalitions: Optimality and stability*, Econometrica: Journal of the Econometric Society (1980), 987–1003.
- Moran Feldman, Liane Lewin-Eytan, and Joseph Seffi Naor, *Hedonic clustering games*, ACM Transactions on Parallel Computing **2** (2015), no. 1, 4.