Sergio Zumpano Arnosti – RA:188959
Resumo do artigo: Best-Response Mechanisms - N. Nisan, M. Schapira, G. Valiant, and A. Zohar.

1. Introdução, motivação e objetivos

O objetivo principal do artigo é identificar classes de jogos para as quais os mecanismos de resposta ótima são compatíveis com incentivo, ou seja, quando todos os outros jogadores estão repetidamente respondendo o ótimo então um jogador é incentivado a fazer o mesmo.

Dentro da teoria dos jogos e da economia, o objetivo básico dos estudos é o equilíbrio, um estado estável no qual nenhum dos jogadores quer desviar. Porém, geralmente estes estudos abstraem o modo como o equilíbrio é alcançado e na maioria dos casos uma verdade satisfatória é que cada jogador toma ações locais e racionais que misteriosamente levam o sistema a atingir um equilíbrio global.

As principais questões que o artigo aborda são:

- O comportamento localmente racional é realmente racional?
- É racional para os jogadores repetidamente responderem o ótimo?
- Pode um jogador mais esperto melhorar ao longo do jogo através dessa otimização míope?

Os autores exploram essas questões em jogos de informação incompleta e não acoplada, onde se observa que repetidamente responder o ótimo traz uma característica interessante: para responder o ótimo cada jogador necessita apenas saber sua própria função de utilidade, uma vez que a sua resposta ótima não depende das funções dos outros jogadores, mas apenas das ações deles. Nesse contexto, a dinâmica de resposta ótima pode ser vista como um protocolo natural onde o compartilhamento limitado e gradual de informações é um esforço para alcançar o equilíbrio. De fato, em muitos contextos do mundo real a interação entre aqueles que tomam decisões com informações incompletas tem a forma de uma dinâmica de resposta ótima. (ex.: roteamento na internet).

2. A configuração

2.1. Jogo base

No *framework* utilizado pelos autores, cada jogador tem uma função de utilidade privada e quando as funções de todos os jogadores são colocadas juntas elas determinam um jogo base de informação completa.

Em um jogo base G de n jogadores (1, …,n), cada jogador i tem um tem um espaço de estratégias S_i, e $S=S_1\times\ldots\times S_n$. Além disso, cada jogador i tem uma função utilidade particular u_i tal que $(u_1,\ldots,u_n)\in U\subseteq U_1\times\ldots\times U_n$ onde $U_i\subseteq R^{|S|}$ é o espaço de estratégias de i.

2.2. Mecanismos de resposta ótima

Foi estudada uma classe de mecanismos nomeados mecanismos de reposta ótima. Os jogadores passam turnos selecionando estratégias, onde a cada passo de tempo t, algum jogador i_t seleciona e anuncia uma estratégia $s_i^t\in S_{i_t}$. A cada turno uma ação disponível para cada jogador em um mecanismo de repetição é sempre escolher a resposta ótima para as respostas mais recentes anunciadas pelos outros jogadores.

Para especificar completamente um mecanismo de resposta ótima é necessário determinar um estado inicial, uma ordem de ativação dos jogadores e pra cada jogador uma regra de desempate entre várias respostas ótimas.

2.3. Regras de desempate

Quando existem múltiplas respostas ótimas, é necessário especificar para cada jogador uma regra de desempate. Esta regra deve depender somente da informação privada do jogador (sua função utilidade). Nos jogos modelados no artigo, as regras fixam para cada jogador i uma ordem $\prec_i$ em S_i. Dessa forma, quando os jogadores possuem múltiplas respostas ótimas eles são instruídos a desempatar de acordo com $\prec_i$.

3. Jogos com mecanismos de resposta ótima compatíveis com incentivo

Através de um jogo simples os autores mostram que nem todo jogo no qual os jogadores respondem repetidamente a resposta ótima é compatível com incentivo. Então, apresentam uma classe de jogos conhecidas pelo termo *"Never-Best-Response-Solvable (NBR-Solvable) with clear outcomes*". Em um jogo *NBR-Solvable* estratégias são iterativamente eliminadas se essas nunca levarão a um resultado ótimo. Essa classe de jogos tem um resultado claro/objetivo *(clear outcome)* se um jogador i considera que não pode fazer melhor do que o resultado do jogo que foi alcançado via a repetição de respostas ótimas, dado que os outros jogadores já eliminaram as estratégias independentemente das ações de i. Apresentada essa classe de jogos, os autores definem um teorema informal.

Teorema (informal)

Defina G como um jogo *NBR-Solvable* com um resultado claro. Então, para todo ponto de início e para toda ordem de ativação de jogadores com pelo menos $T=\sum_i |S_i|-n$ turnos, temos:

- A dinâmica de repetidamente responder o ótimo converge para um equilíbrio puro de Nash s^* de G;
- A dinâmica de repetidamente responder o ótimo é compatível com incentivo.

4. Formalização e modelagem

O artigo organiza e formaliza todas as definições necessárias para a modelagem e apresenta um teorema geral para o mecanismo compatível com incentivo.

As definições apresentadas e que podem ser consultadas no artigo [1] são para:

- Regras de desempate *(tie-breaking rules)*
- Estratégias *NBR*
- Jogos *NBR-Solvable*
- Parâmetros para a menor sequência de eliminação
- Perfil global ótimo
- Resultados claros *(clear outcomes).*

Teorema 1 (mecanismos compatíveis com incentivo)

Defina G como um jogo *NBR-Solvable* com um resultado claro s^* sujeito a $\prec_1, \ldots, \prec_n$. Defina M como um mecanismo de resposta ótima para G. Então, para todo ponto de início e para toda ordem de ativação dos jogadores com pelo menos T turnos, onde um turno é uma sequencia consecutiva de passos no tempo na qual cada jogador está "ativo" ao menos uma vez, então temos:

1. M converge para s^*;
2. M é compatível com incentivo.

A prova deste teorema é feita primeiramente para a convergência, onde os autores mostram que uma estratégia *NBR* nunca será escolhida por um jogador durante os turnos e no subjogo G_0 ela será eliminada, fazendo com que esse subjogo seja equivalente ao próximo subjogo, G_1. E seguindo a ordem de eliminação de estratégias *NBR*, o jogo convergirá para um último jogo G_r onde a única estratégia será s^* que é o único equilíbrio puro de Nash sujeito às regras de desempate de G.

Depois de provada a convergência, os autores provam a parte de compatibilidade com incentivo. Por contradição é suposto que exista um jogador que desvie da dinâmica de sempre responder o ótimo, dessa forma seguindo a mesma linha da prova da convergência é mostrado que o último subjogo obtido pela eliminação prévia de todas as estratégias *NBR* terá apenas um perfil ótimo global s^* independente das ações do jogador que desviou da dinâmica e dessa forma esse jogador não tem como fazer melhor do que s^* e por isso é incentivado a sempre responder o ótimo.

5. Exemplos de jogos

Formalizadas as definições e a prova do teorema principal de mecanismos de resposta ótima compatíveis com incentivo, os autores analisam quatro ambientes que podem ser modelados em jogos *NBR-Solvable* com respostas claras, sujeitos a regras de desempate e compatíveis com incentivo. Apenas três dos quatro cenários estudados, considerados os mais relevantes, serão apresentados nesse resumo.

5.1. Stable-roommates

Nesse jogo existem como jogadores n estudantes e cada um possui um *ranking* estrito e privado dos outros, os participantes preferem ser pareados a não ser. Cada estudante i possui um valor $\alpha_i(j)$ que é o *rank* que i coloca para j. O estudante i possui uma utilidade $u_i(s)=\alpha_i(j) \Leftrightarrow s_i=j$ e $\nexists k \neq i$ tal que $s_k=j$ e $\alpha_j(k)>\alpha_j(i)$, caso contrário $u_i(s)=0.$

Em geral, dentro desse contexto não há garantias que exista um pareamento estável, porém os autores analisam casos nos quais o pareamento estável é garantido e alcançado através de incentivo. São eles:

- *Intern-hospital matchings*: Os jogadores são divididos em dois conjuntos disjuntos (residentes e hospitais). Cada um dos hospitais possui o mesmo *ranking* para os residentes.
- *Correlated markets*: Os "estudantes" são vértices em um grafo, toda aresta tem um único "peso" e quanto mais pesada é a aresta conectando estudantes, maior é um *rank* dado por aquele estudante ao outro conectado à ele.

Teorema 2 (stable-roommates games)

Para todo jogo G e para ambos os casos *Intern-hospital e Correlated markets* temos que:

- G é *NBR-Solvable* e o único equilíbrio puro de Nash é um pareamento estável;
- $e_G \leqslant n$ (o número de turnos é menor ou igual ao número de jogadores).

A prova do teorema considerando que o jogo é livre de ciclo *(cycle-free)* de tamanho maior que 2 segue uma sequencia de eliminação de estratégias, uma vez que esta é necessária para que o jogo seja *NBR-Solvable*. Então seleciona-se um estudante arbitrário r_1 e constrói-se uma sequência r_1, r_2, …, na qual o estudante r_{i+1} é o estudante que r_i prefere. Sendo o número de estudantes finito, a sequência deverá repetir em algum momento e levando em consideração que o jogo é livre de ciclo, o ciclo deve ser de tamanho 2 e dessa maneira foram localizados dois estudante que preferem um ao outro. Logo, podemos eliminar para esses dois estudantes as estratégias que eles tem de propor para outros, uma vez que a utilidade máxima foi obtida na proposta em que eles se preferem.

O artigo também mostra que os dois casos específicos estudados não induzem um ciclo no grafo de pareamento e dessa forma o mecanismo para *stable-roommates* é um mecanismo de resposta ótima e pelo teorema 1 o mecanismo é compatível com incentivo.

5.2. Roteamento na internet

No roteamento entre os sistemas autônomos da internet, o BGP é um protocolo que estabelece as rotas entre eles.

Nesse jogo a rede é um grafo não direcionado G= (V, E), onde V é o conjunto de nós fonte (1, …,n) e tem um único destino d. Cada um dos nós fonte tem um *ranking* estrito de rotas (sem ciclos) que levam até o destino d. Um nó fonte i possui um conjunto de estratégias S_i composto pelas arestas de saída e também um vetor $\vec{f}=(f_1,\dots,f_n)$ que representa as decisões de encaminhamento dos nós fonte. Sua utilidade $u_i(\vec{f}=(f_1,\dots,f_n))$ é igual ao *rank* que i coloca para a rota até d caso ela exista, e $u_i(\vec{f})=0$ caso contrário.

No BGP, cada fonte repetidamente examina os anúncios mais recentes feitos pelos seus vizinhos e encaminha as suas informações através da rota preferida, depois anuncia a sua escolha aos seus vizinhos.

Segundo um teorema apresentado no trabalho de Levin *et al* [2] o BGP é compatível com incentivo se não existir uma condição de *Dispute Wheel*, onde um nó prefira usar um outro nó pivô ao invés de utilizar sua rota direta para chegar ao destino d.

Teorema 3 (BGP games)

Sendo G um jogo BGP, temos que:

- G é *NBR-Solvable* e o único equilíbrio é uma árvore estável de roteamento;
- $e_G \leqslant n$.

A prova do teorema primeiro define a sequencia de eliminação de estratégias *NBR* de forma que podemos sempre encontrar um nó que pode garantir a sua rota mais preferida até o destino d e dessa forma pode-se eliminar todas as outras ações de rotas.

Inicia-se o jogo com um nó arbitrário a_0 com pelo menos duas ações, então é definida a rota R_0 como sendo a rota preferida desse nó até o destino d. Escolhe-se então um vértice a_1 em R_0 que está mais próximo do destino d tal que ele prefira a rota R_1 que leva a d. Faz-se isso então para os próximos nós da rota até não existam mais ações a serem tomadas e se atinja o nó destino.

Dessa forma, foi definida uma sequencia de eliminação de estratégias *NBR* e também foi obtida uma árvore estável de roteamento de um nó fonte até o destino d. Considerando também que não existe a condição de *Dispute Wheel*, então o jogo BGP é compatível com incentivo.

5.3. Controle de congestionamento

O controle de congestionamento na internet é feito através de uma combinação de técnicas do protocolo TCP *(Transmission Control Protocol),* nas quais há o ajuste das taxas de transmissão e políticas de gerenciamento de filas.

No artigo, o jogo do protocolo TCP modelado é uma rede formada por um grafo G= (V, E) não direcionado onde $c(e)$ especifica a capacidade de uma aresta $e \in E$. Também existem n pares fonte-destino (α_i, β_i) que desejam transmitir ao longo de uma rota R_i em G. Um nó α_i pode selecionar um taxa de transmissão no intervalo $[0, M_i]$ e sua utilidade $u_i(\vec{r})$ é a vazão alcançada em um fluxo em equilíbrio onde $\vec{r}$ é o vetor das taxas de transmissão utilizadas pelos nós fonte.

O jogo modelado é baseado no estudo de Godfrey *et al* [3] que analisa os cenários inspirados em TCP compatíveis com incentivo e considera um protocolo chamado *Probing-Increase-Educate-Decrease* (PIED) provado ser compatível com incentivo.

Dois esquemas de alocação são considerados:

- *Strict-Priority-Queuing* (SPQ): cada aresta possui um *ranking* dos nós fontes, e o nó fonte com o maior *rank* tem o seu fluxo inteiro passado pela aresta até um máximo $c(e)$; e a capacidade da aresta que sobra é oferecida ao segundo melhor classificado.
- *Weighted-Fair-Queueing* (WFQ): cada nó fonte α_i tem um peso $w_i(e)$ na aresta e . A capacidade alocada para α_i será $\frac{w_i}{\Sigma_j w_j} c(e)$. O caso em que $\forall e \in E$, $\forall i \in [n], w_i(e) = 1$ é chamado *"fair queuing"* (FQ).

Teorema 4 (TCP games)

Sendo G um jogo TCP para o qual as arestas utilizam SPQ com prioridades coordenadas ou WFQ com pesos coordenados, temos que:

- G é *NBR-Solvable* e tem um único equilíbrio puro de Nash que é um padrão de fluxo estável;
- $e_G \leqslant n$.

A prova para o teorema é feita apenas para o caso WFQ com $w_i(e) = 1$, *"fair queuing"*. Então para cada aresta e teremos que cada fluxo poderá utilizar uma capacidade $\beta_e = c_e / k_e$, onde c_e é a capacidade da aresta e k_e é a quantidade de fluxos que utilizam aquela. Dessa forma pode-se construir uma sequencia de eliminação das estratégias *NBR*.

Sendo uma aresta e^* com uma razão β mínima, então cada fluxo nessa aresta terá um tráfego garantido de exatamente β_{e^*} e pelo menos essa mesma quantidade nas outras arestas. Dessa forma, é possível eliminar todas as ações de se transmitir menos que β_{e^*} . Se um jogador i elimina essas ações por último entre os jogadores então ele fará isso em um jogo no qual o perfil final é o ótimo para ele. Com isso vemos que o jogo tem uma reposta clara e o teorema 1 implica que ele é compatível com incentivo.

6. Conclusões

Através dos casos analisados foi possível explorar quando uma dinâmica de racionalização local é também uma racionalização global.

Os resultados ao longo do artigo incentivam os estudiosos da área de redes a pensarem em novas estruturas para os protocolos da internet e proveem novas visões para projetar futuros protocolos e mecanismos. Além disso, o artigo se mostra interessante por apresentar, mesmo que em condições específicas, que cenários reais com mecanismos de repetidamente responder o ótimo podem ser compatíveis com incentivo.

7. Referências

[1] N. Nisan, M. Schapira, G. Valiant, A. Zohar. “Best-Response Mechanisms”, Innovations in Computer Science - ICS 2011, pages 155-165, 2011.

[2] H. Levin, M. Schapira, and A. Zohar. “Interdomain routing and games”. In STOC, pages 57-66, 2008.

[3] P. B. Godfrey, M. Schapira, A. Zohar, and S. Shenker. “Incentive compatibility and dynamics of congestion control”. SIGMETRICS Perform. Eval. Rev., 38(1):95-106, 2010.