VE444: Networks

Yifei Zhu, assistant professor
University of Michigan-Shanghai Jiao Tong University

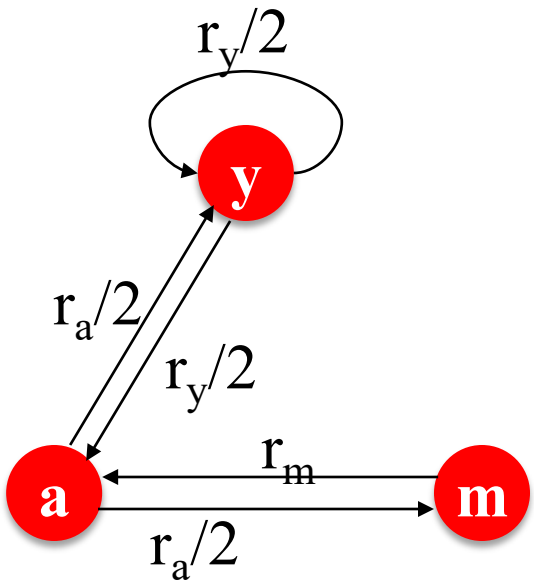
PageRank: The "Flow" Model

- **A page is important if it is pointed to by other important pages**
- **Define a "rank" r_j for node j**

$$r_j = \sum_{i \to j} \frac{r_i}{d_i}$$

d_i … out-degree of node i

The web in 1839

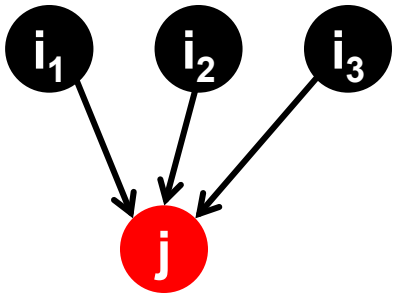


"Flow" equations:

$$r_y = r_y/2 + r_a/2$$
$$r_a = r_y/2 + r_m$$
$$r_m = r_a/2$$

You might wonder: Let's just use Gaussian elimination to solve this system of linear equations. Bad idea!

Random walk



- **Where is the surfer at time $t+1$?**
  - Follow a link uniformly at random

$$\boldsymbol{p(t+1) = M \cdot p(t)}$$

- Suppose the random walk reaches a state

$$\boldsymbol{p(t+1) = M \cdot p(t) = p(t)}$$

  then $\boldsymbol{p(t)}$ is **stationary distribution** of a random walk

- **Our original rank vector $\boldsymbol{r}$** satisfies $\boldsymbol{r = M \cdot r}$
  - **So, r is a stationary distribution for the random walk**

Solution: Random Teleports

- **Google's solution that does it all:**
  At each step, random surfer has two options:
  - With probability β, follow a link at random
  - With probability $1\text{-}\beta$, jump to some random page

- **PageRank equation** [Brin-Page, 98]

$$r_j = \sum_{i \to j} \beta \frac{r_i}{d_i} + (1 - \beta)\frac{1}{N}$$

d_i ... out-degree of node i

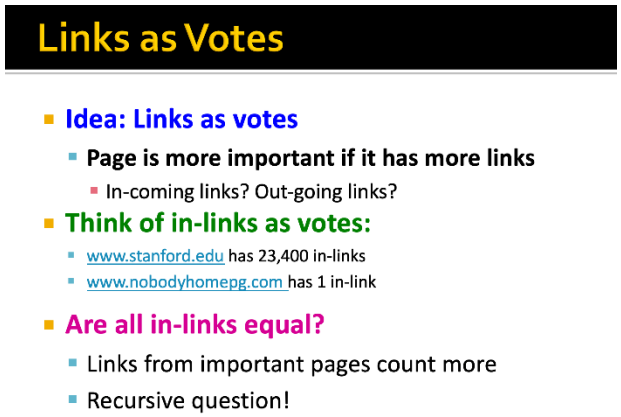
This formulation assumes that $\boldsymbol{M}$ has no dead ends. We can either preprocess matrix $\boldsymbol{M}$ to remove all dead ends or explicitly follow random teleport links with probability 1.0 from dead-ends.

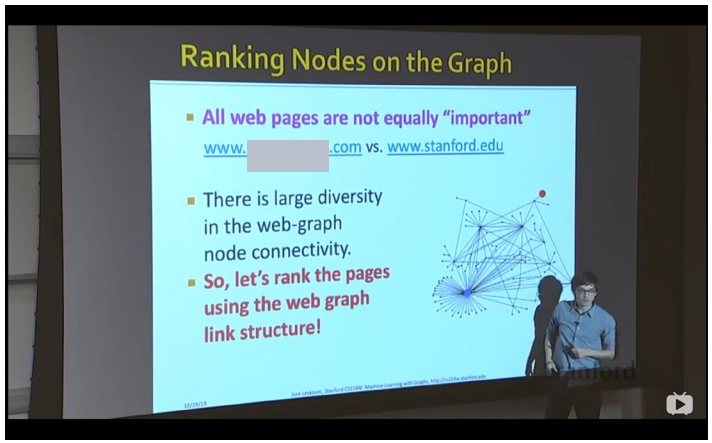
Link analysis in Modern Web search

- **Combining Links, text, and usage data**
  - **"I am a student at SJTI-JI". (Anchor text)**
  - **"For more information, click here"**

Link analysis in Modern Web search

- Current ranking functions are constantly evolving
- Always expect the world to react to what you do (Game theoretic view)
  - Search Engine Optimization/Marketing(SEO, SEM)
  - An example last week





https://www.bilibili.com/video/BV1jE41177A4?p=13

Sponsored Search Markets

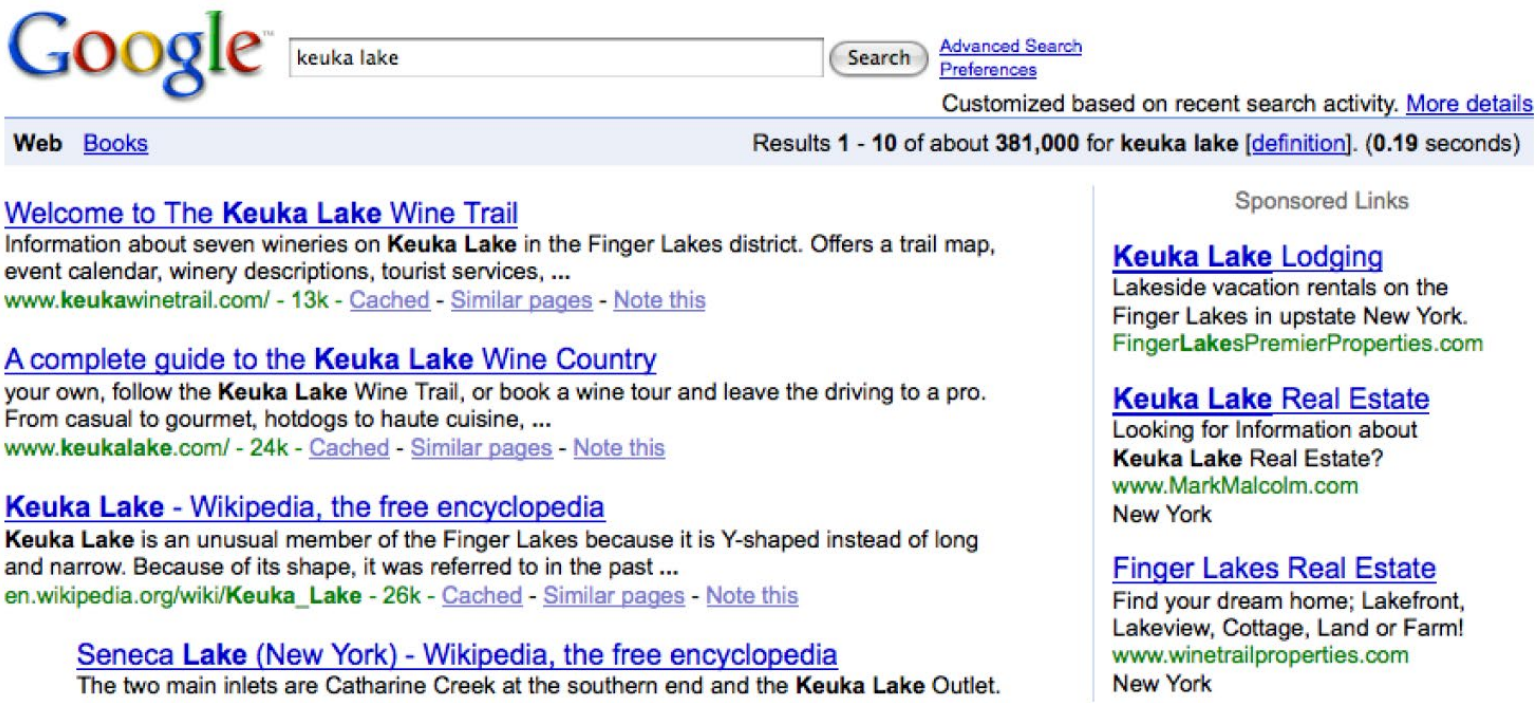
Google revenue breakdown

- In 2019, Google made $161.85 billion in revenue
- Out of this amount, $134.81 billion from advertising!

70.9%

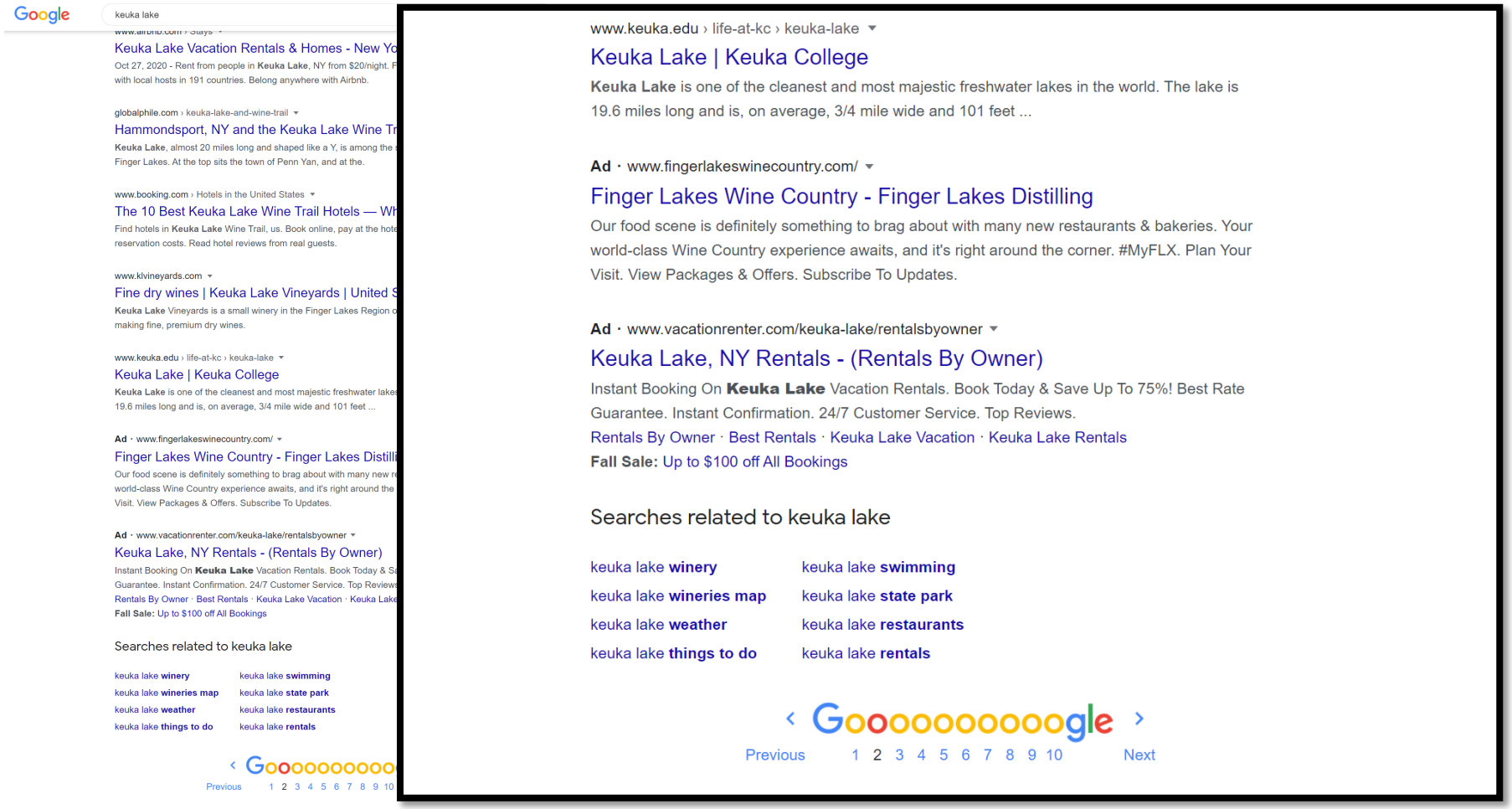
Sponsored search market

- The search advertising market



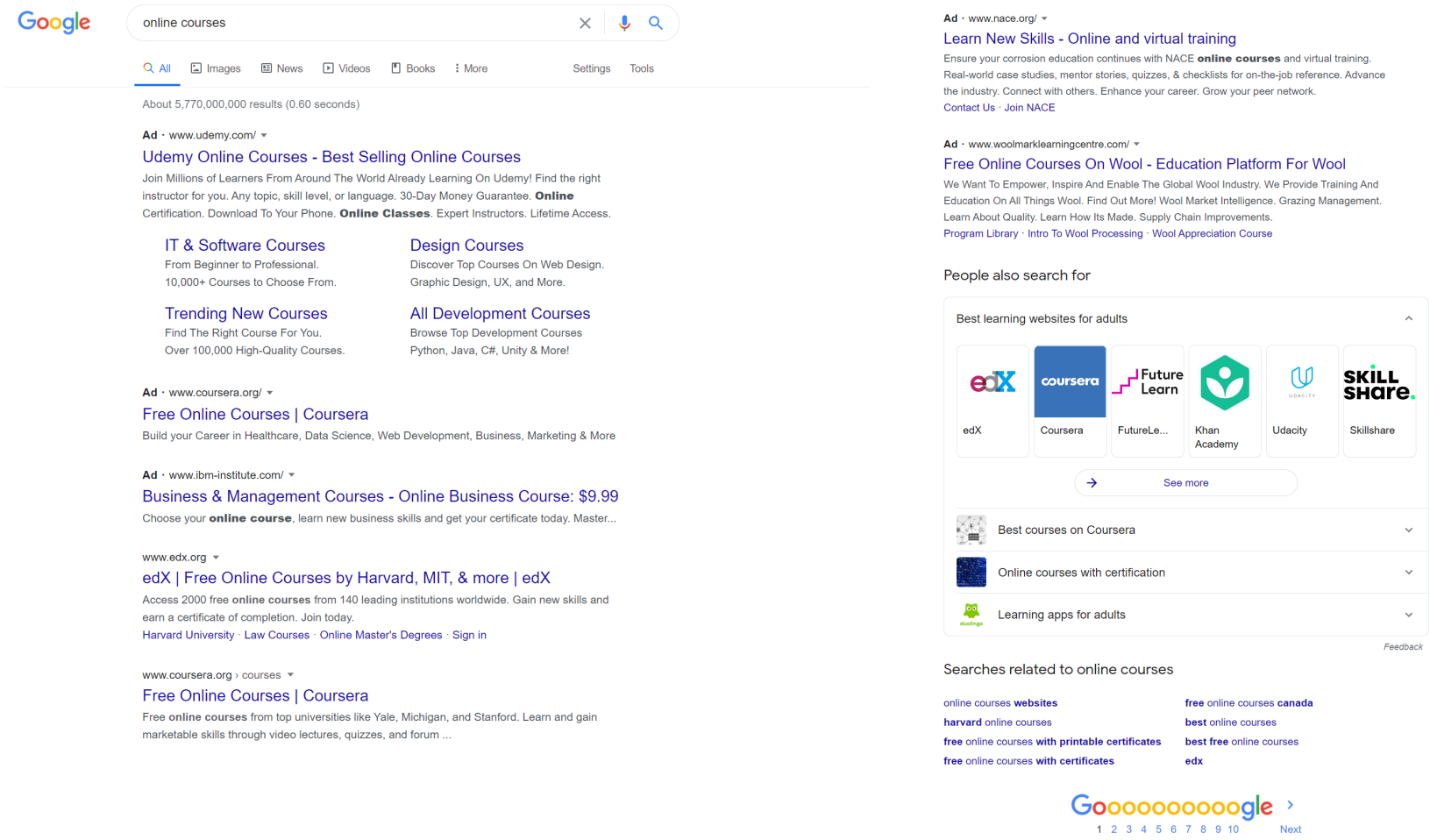
Sponsored search market

- The search advertising market



Sponsored search market

- The search advertising market



Sponsored search market

- We have studied auction
- We have studied matching
- We study them in separation!

- Sponsored search market is a combination of matching and auction!

Three problems

- Question 1: how should advertisers pay for the ads?
- Question 2: how should websites allocate the ads slots to the advertisers?
- Question 3: how should websites set the prices on the ad slots

- Sponsored search market is a combination of matching and auction!

How to sale ads in search engines

Keyword-base advertisement

- Targeted advertisement, as compared to traditional forced advertisement
- Low cost, especially for less frequently searched keywords
- Easy to evaluate the effectiveness, i.e., the number of users clicked

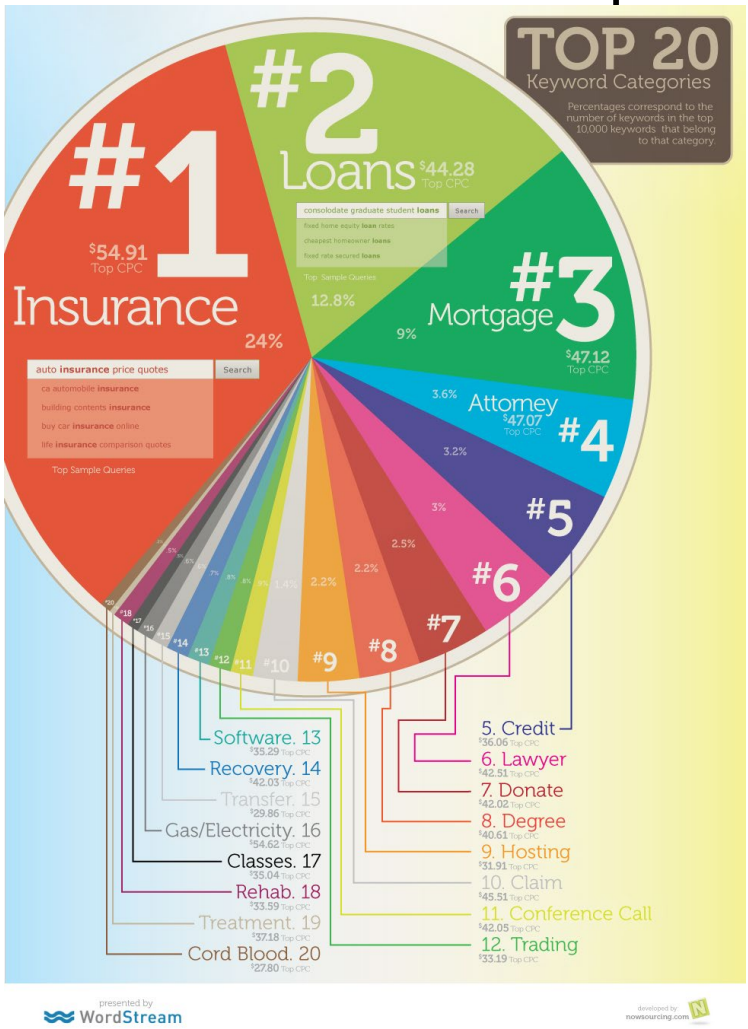
Question 1: how should advertisers pay for the ads?

Payment methods by the websites

- By showing:
  - CPM (cost per mille/cost per thousand impressions)
- By action:
  - CPA (cost per action)
  - CPC (cost per click)
- By sale:
  - CPO (cost per order)

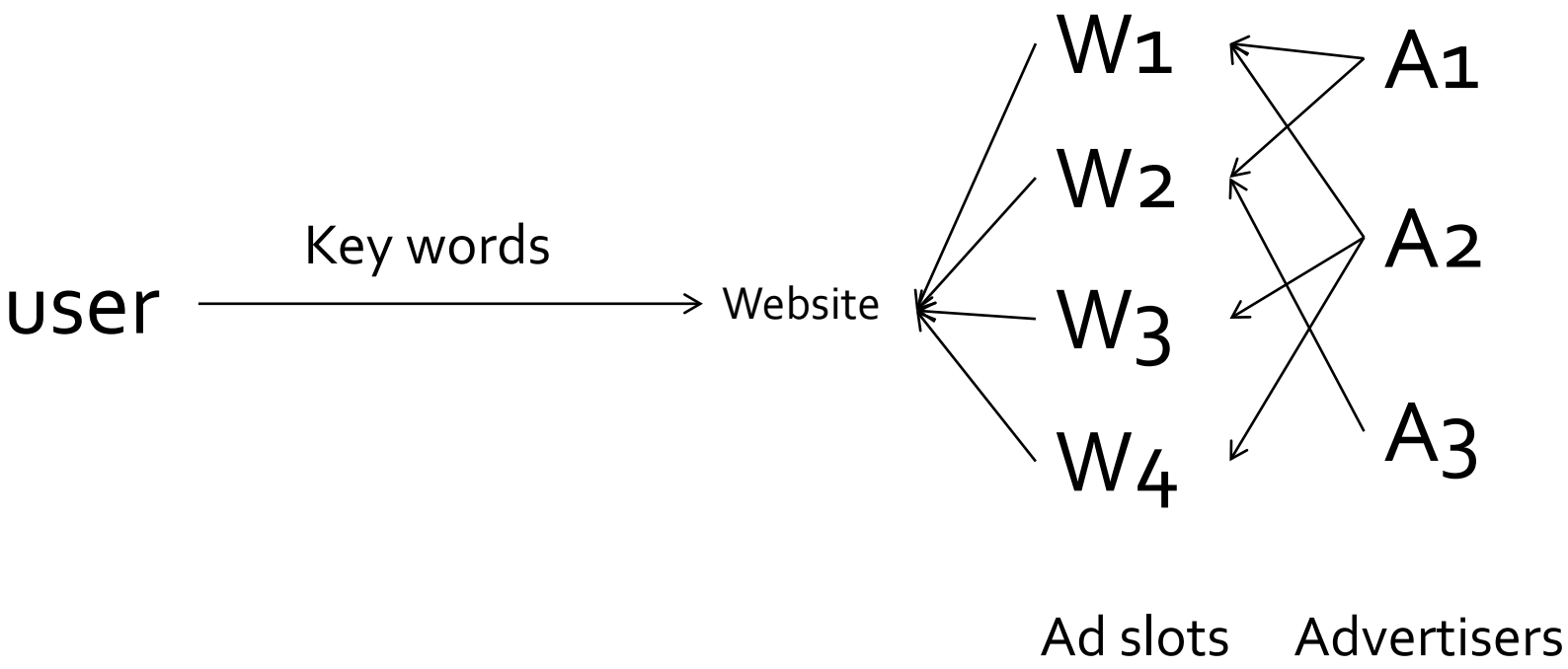
Payment methods by the websites

- Different key words have different prices



Keyword-base advertisement

Question 2: how should websites allocate the ads slots to the advertiser?



Modeling

- Can we model it into the matching problem we have studied before?

- From a real world scenario → problem formulation (recall the matching market that we have learned before)

Some basic concepts

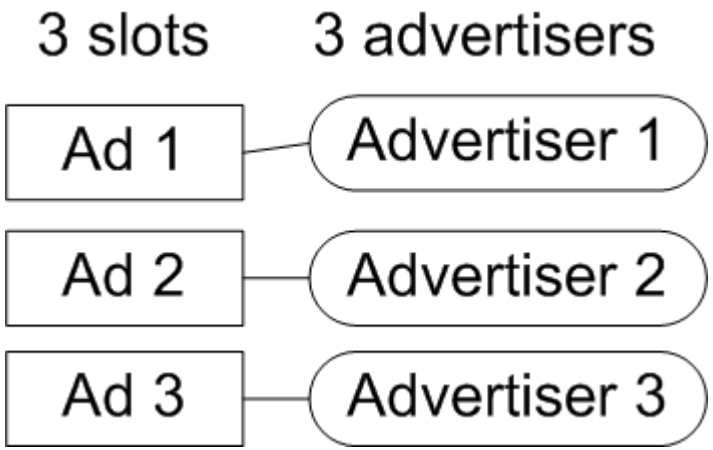
- Advertising slots
- Clickthrough rate: the expected clicks per hour on an advertising slot
- Advertiser's revenue per click: the expected revenue of every click

Some basic concepts

- Advertising slots
- Clickthrough rate: the expected clicks per hour on an advertising slot
- Advertiser's revenue per click: the expected revenue of every click

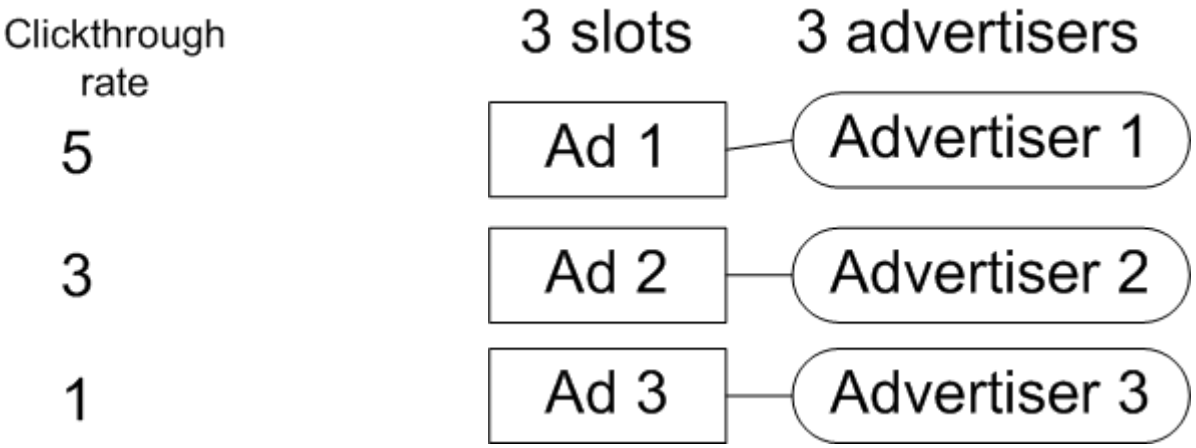
- Advertiser's valuation (per unit time): clickthrough rate * revenue per click
- Advertiser's payoff: expected revenue - price

An example



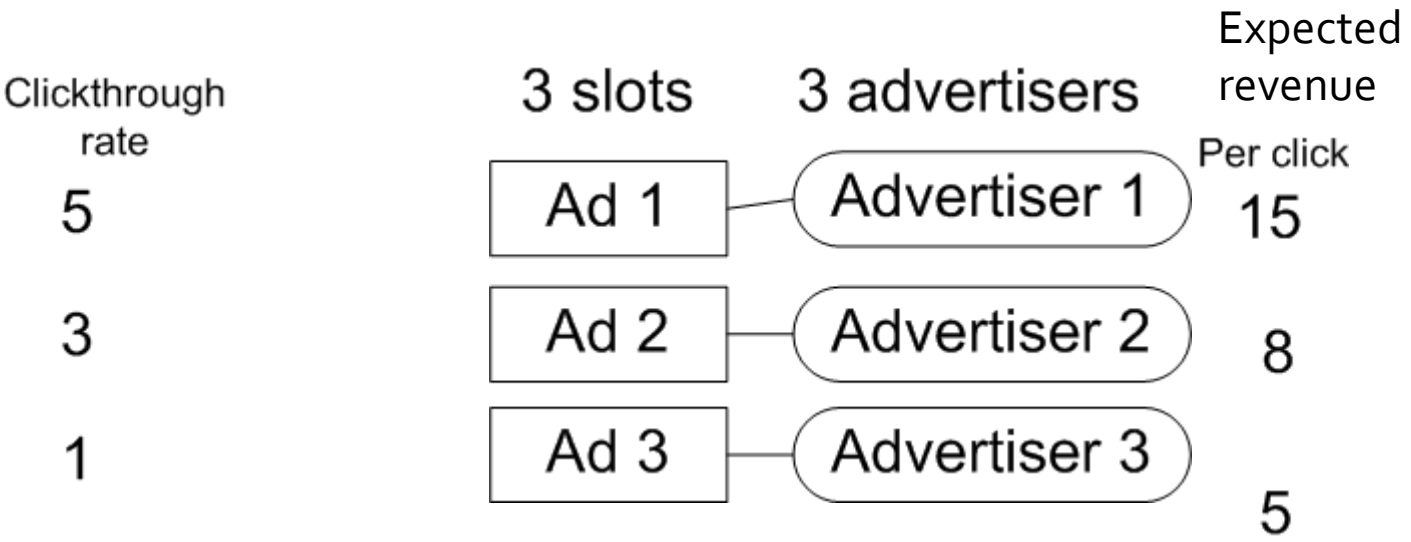
- 3 slots and 3 advertisers

An example



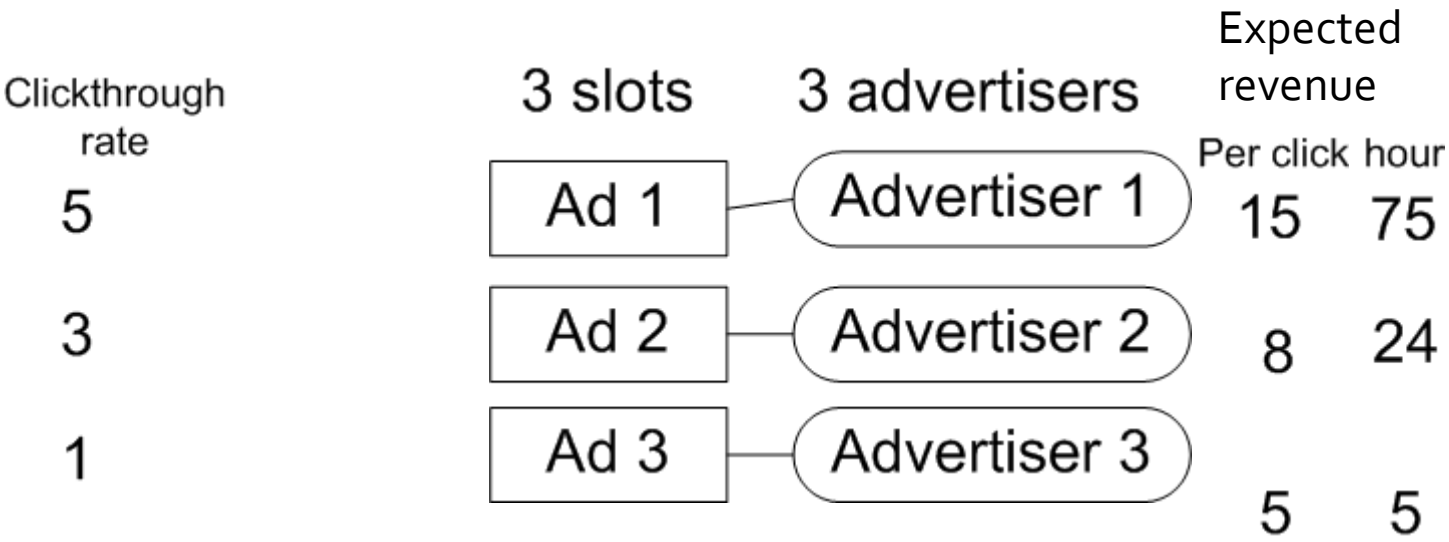
- 3 slots and 3 advertisers
- Clickthrough rate

An example



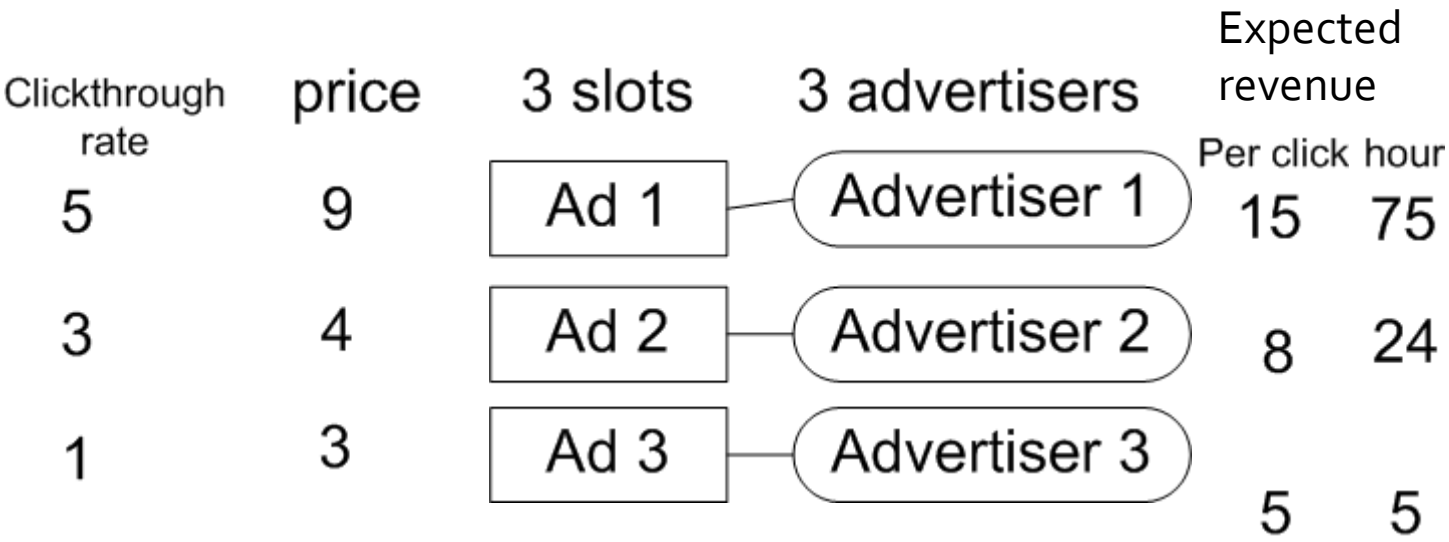
- 3 slots and 3 advertisers
- Clickthrough rate, expected revenue per click

An example



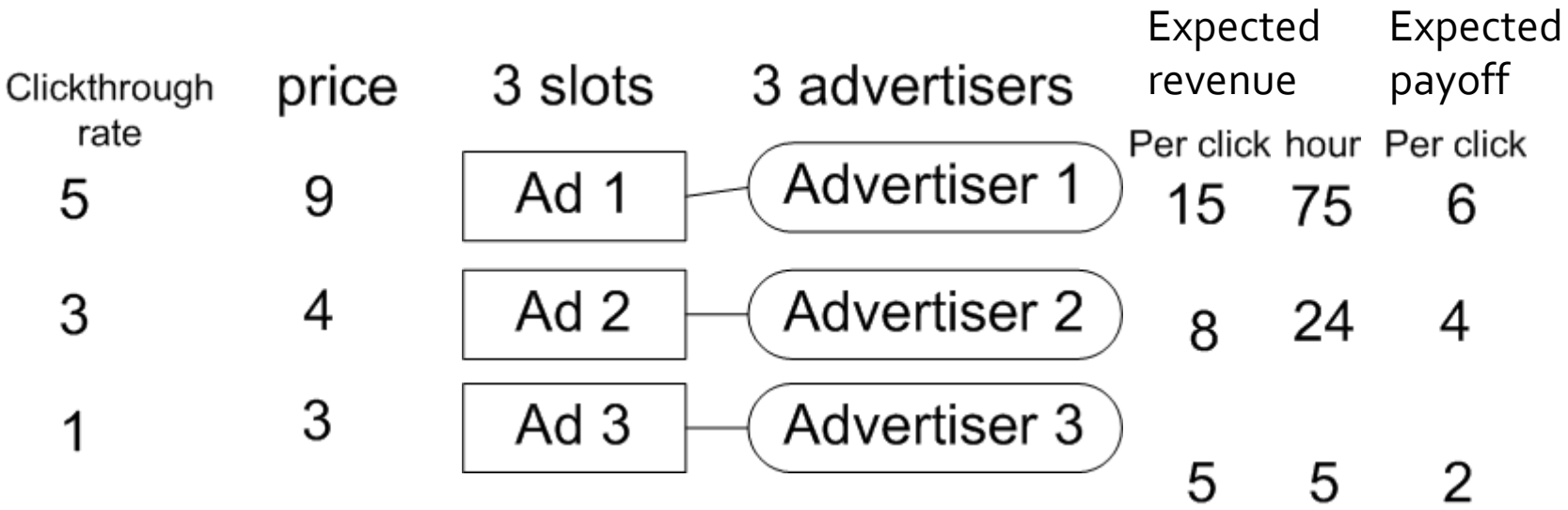
- 3 slots and 3 advertisers
- Clickthrough rate, expected revenue per click
- Valuation = Expected revenue can be calculated = clickthrough rate * expected revenue per click

An example



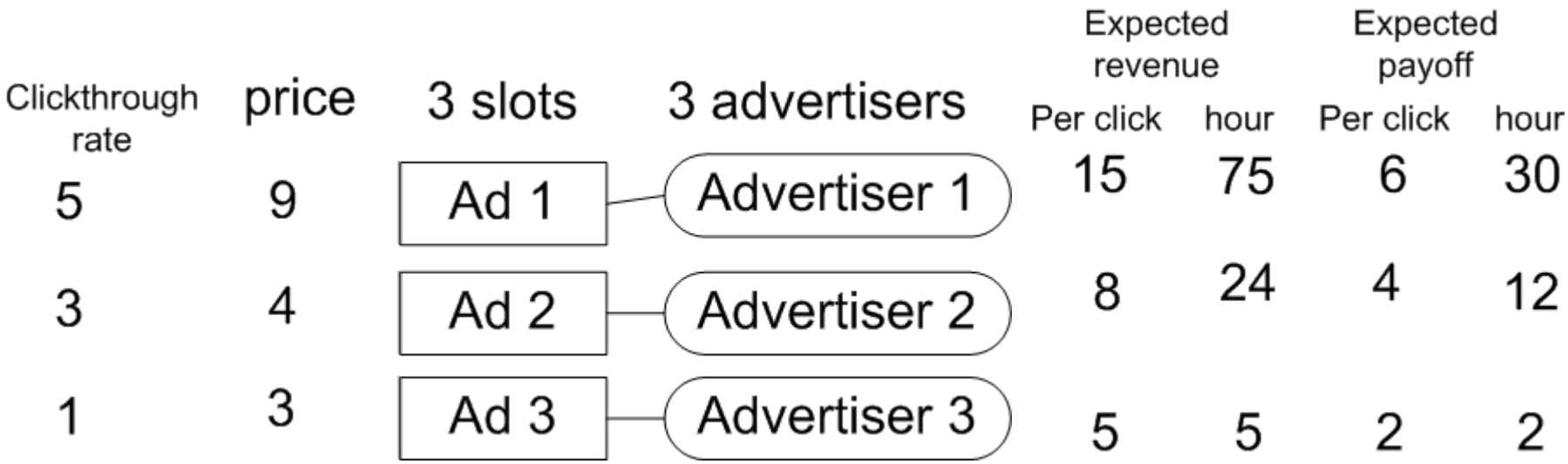
- 3 slots and 3 advertisers
- Clickthrough rate, expected revenue per click
- Valuation = Expected revenue can be calculated = clickthrough rate * expected revenue per click
- Assume we know the price

An example



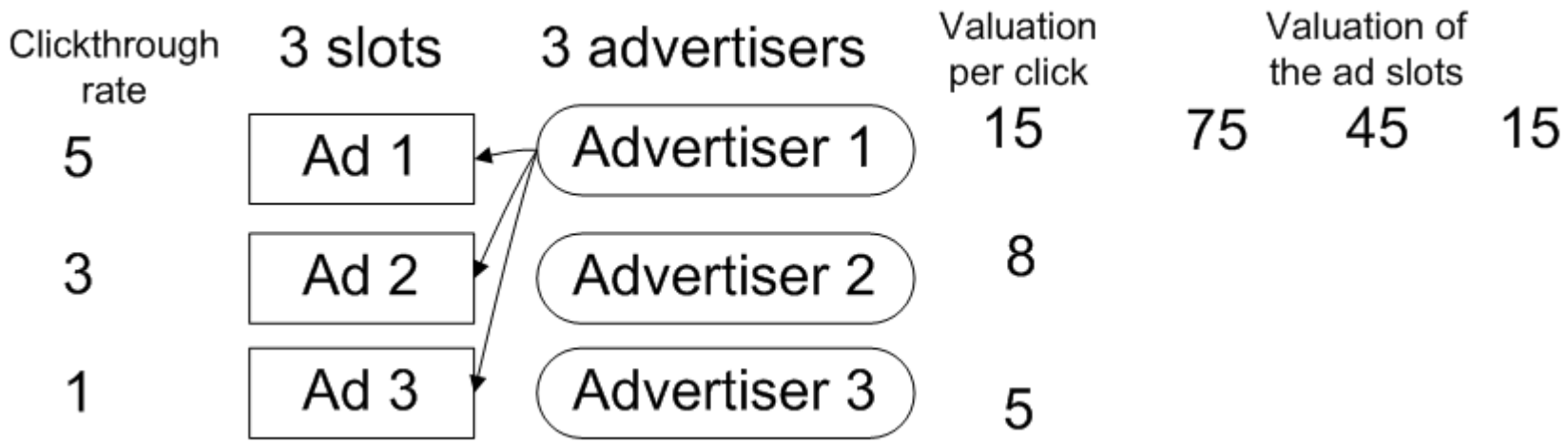
- 3 slots and 3 advertisers
- Clickthrough rate, expected revenue per click
- Valuation = Expected revenue can be calculated = clickthrough rate * expected revenue per click
- Assume we know the price, expected payoff per click

An example



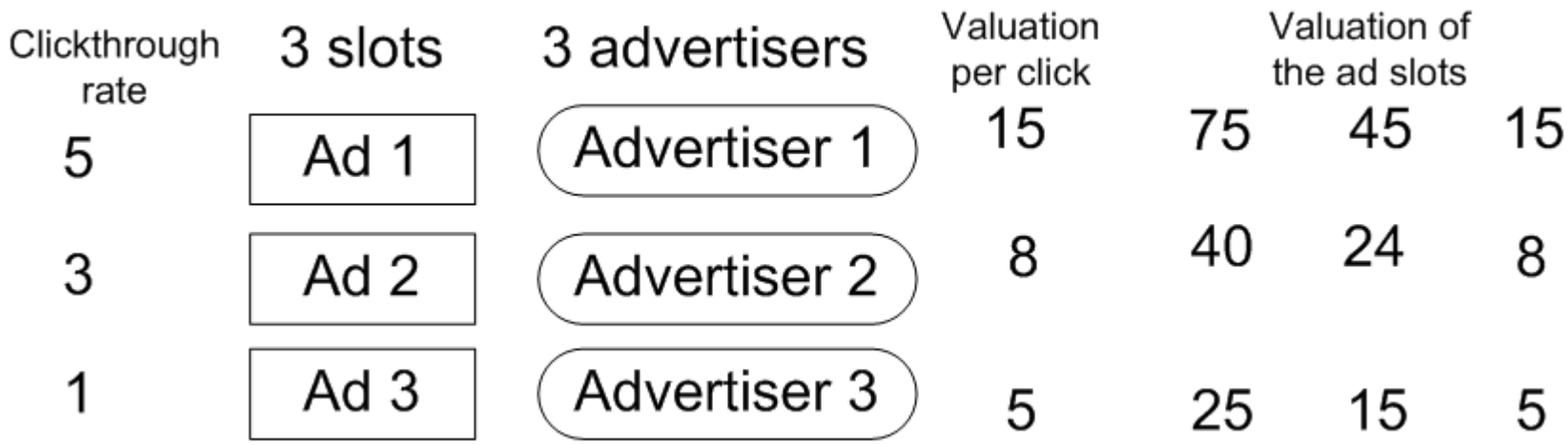
- Advertiser's valuation (per unit time): clickthrough rate * expected revenue per click
- Advertiser's payoff: (expected revenue – price) * clickthrough rate

An example



- Different slots have different clickthrough rates, so their values are different

An example



- Different slots have different clickthrough rates, so their values are different
- Different advertiser's valuation on per click is different, so the valuation of each ad slot is also different

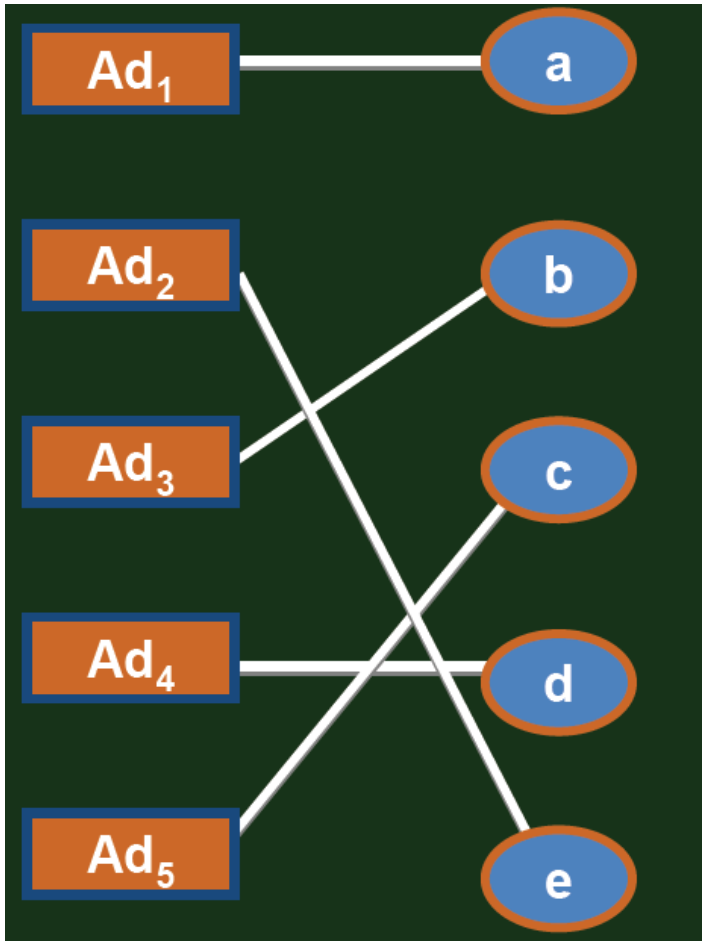
An example

- Seems familiar?
- We have formulated the problem as a matching market

Pricing the ad slots or the clicks

Question 3: how should websites set the prices on the ad slots

- Matching market: market clearing prices
- Vickrey-Clarke-Groves (VCG) mechanism
- Generalized Second-Price Auction (GSP)



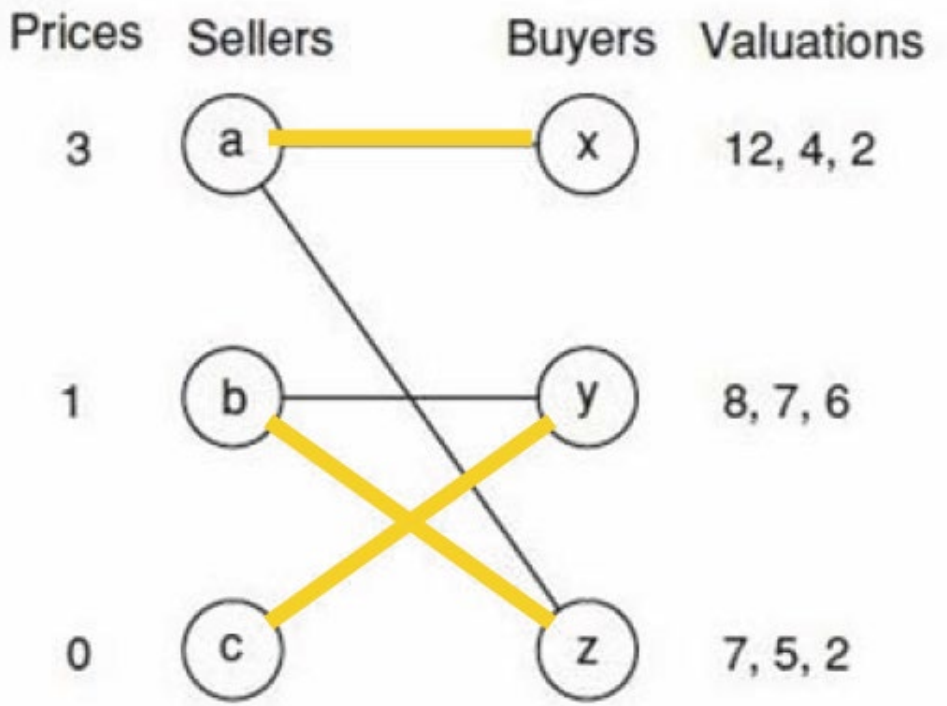
A summary

- General backgrounds on online ads
- Some general concepts
  - Ad slots, clickthrough rate, advertiser's revenue, advertiser's valuation, advertiser's payoff, etc
- The advertisement market
  - Payment methods
  - Ad slots allocation
  - Pricing the ad slots

Price setting:
Matching of advertiser and ad slots

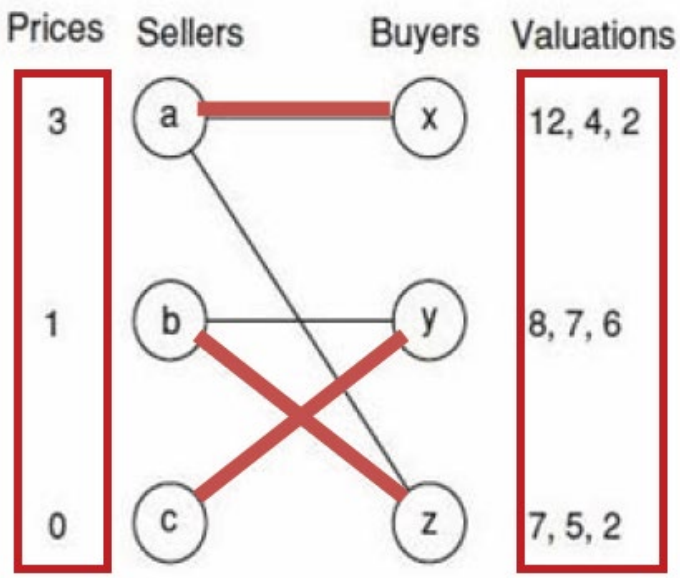
Matching market

- Matching market: market clearing prices



Recall some basic elements

- A set of sellers and buyers
- Every buyer has a valuation on every object from the seller
- Every seller has a price for each object
- Every buyer try to maximize his payoff
- If the preferred sellers graph has a perfect match, then this is an optimal matching, the price in optimal matching is the market-clearing price



Preferred sellers graph

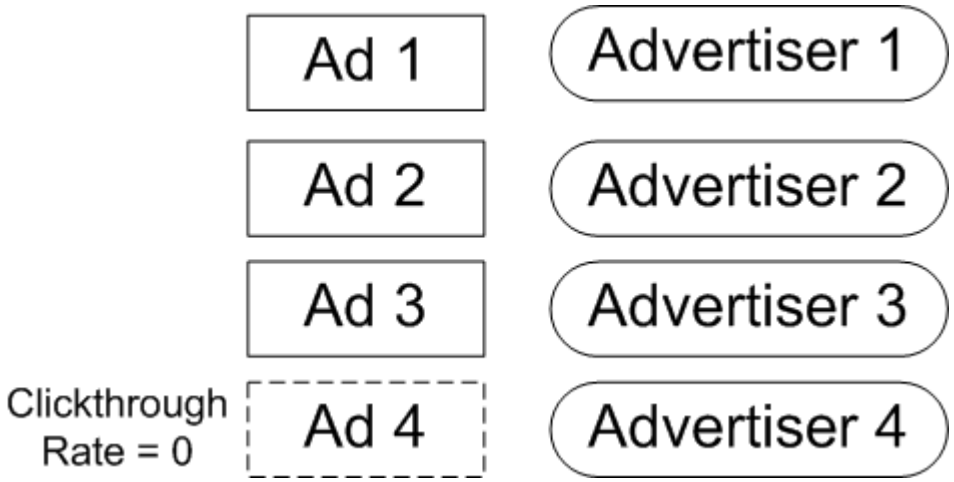
The good property of market-clearing prices

- Market-clearing prices satisfy: optimal matching, maximize the total valuations, maximize the total payoff
- If we know the valuations of the advertisers, every seller uses the market-clearing prices for each slot

The good property of market-clearing prices

- If there are more advertisers than ad slots, we can simply add one slot with clickthrough rate = 0

- 4 advertisers competing for 3 ad slots



The limitation of market matching

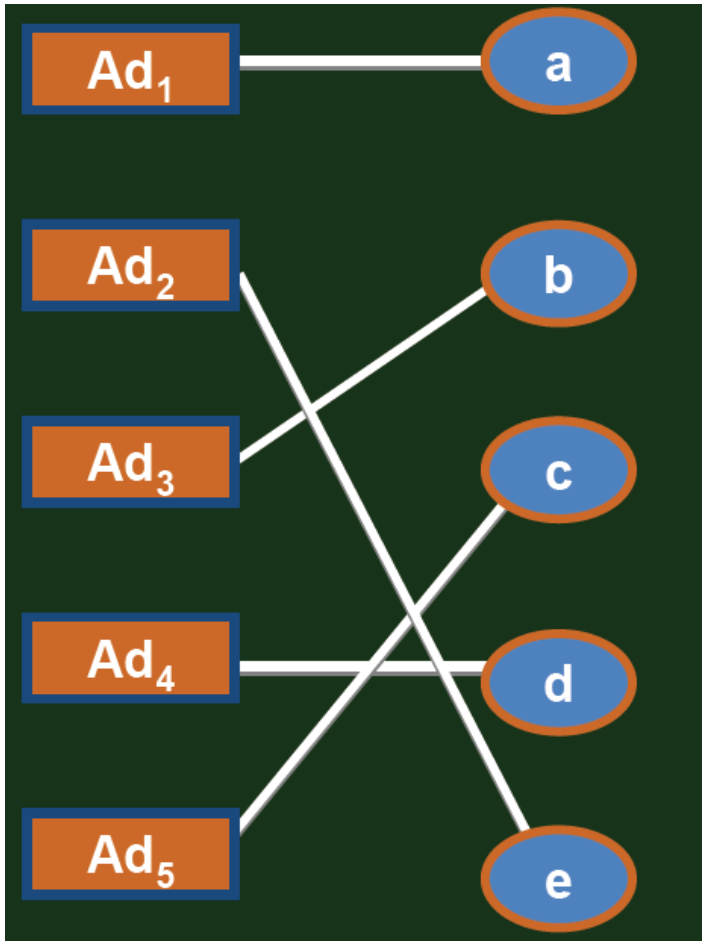
- We have an assumption: every advertiser needs to announce their valuations on the ad slots
- This is not realistic (this can be commercial secrets)

Is there anything else we can do?

Review: pricing the ad slots or the clicks

Question 3: how should websites set the prices on the ad slots

- Matching market: market clearing prices
- Vickrey-Clarke-Groves (VCG) mechanism
- Generalized Second-Price Auction (GSP)



Vickrey-Clarke-Groves (VCG) mechanism

VCG mechanism

- When the advertiser cannot articulate his valuations on the ad slots, e.g., because of secrecy
- Encouraging Truthful Bidding in Matching Markets: The VCG Principle
- This is primarily an auction,
- i.e., if you don't want to say, let's bid

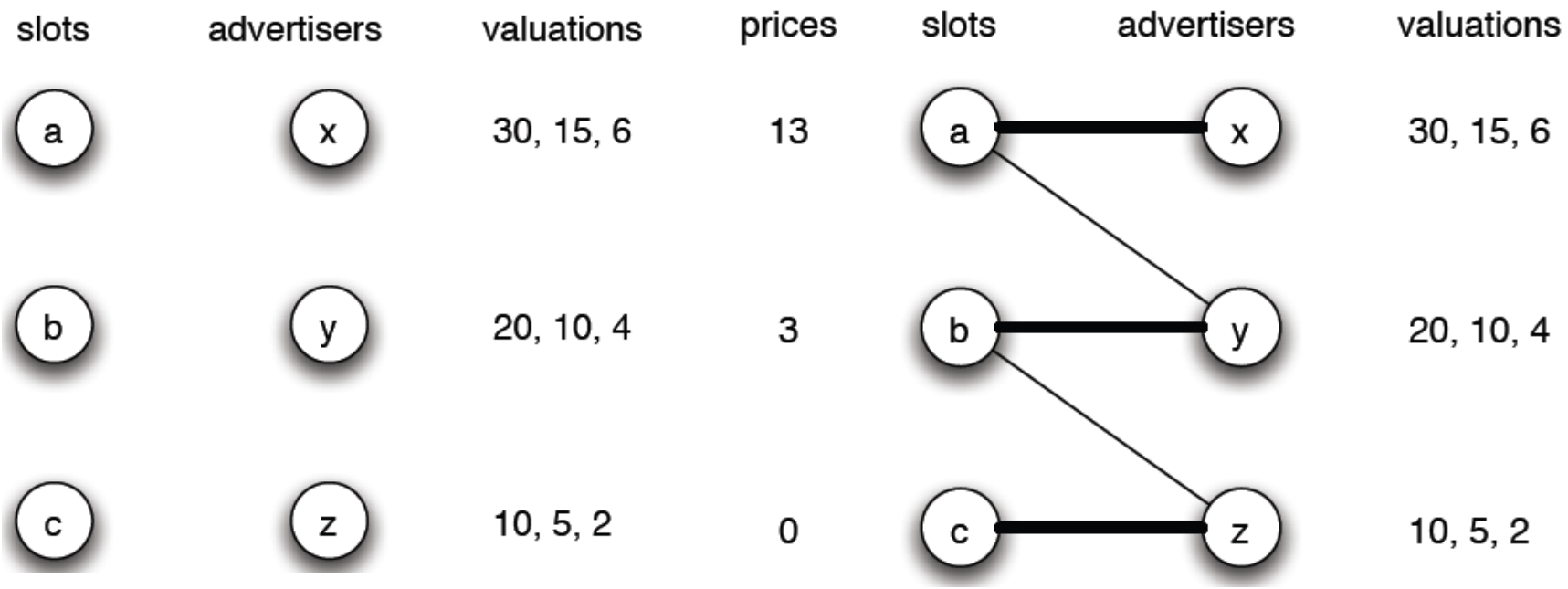
Which auction format to choose?

- First price auction?
  - There are many disadvantages for first price auction
- Can we do second-price auction?
  - We know that for a single item, second-price auction has very good property: truthful bid
- VCG is an extension of single item second-price auction to multiple items

Review: Second-price sealed-bid auctions

- An object
  - Different people may have different values for it, v1, v2, …, vk. These are true value/intrinsic value, i.e., player *i* will pay at most *vi*.
  - Players don't know other's true values
  - Every one has a bid, assume b1 > b2 > … > bk
  - By the rule of second-price auction, the payoff of the highest bidder is v1 – b2 and others are 0; here, b1 is the highest bid, so v1 is the true value
- Truthful bid is a dominant strategy
- Harm due to the existence of the winner

An example



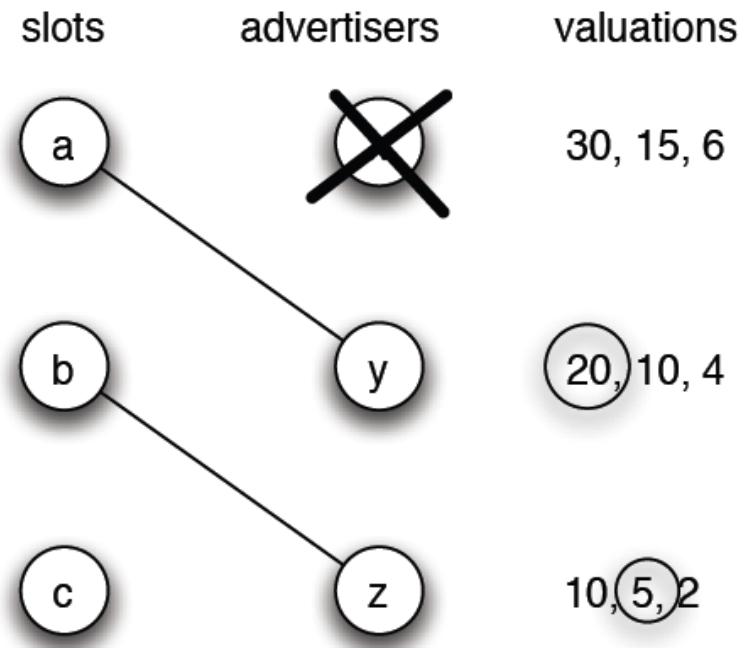
3 slots a, b, c, 3 advertisers x, y z, the valuations are as shown. Using market matching, we can see that the optimal matching is a-x, b-y, c-z, with a total valuation of 42

An example

If x weren't there, y would do better by 20-10=10, and z would do better by 5-2=3

The total harm x caused is 13

This is the price x should pay



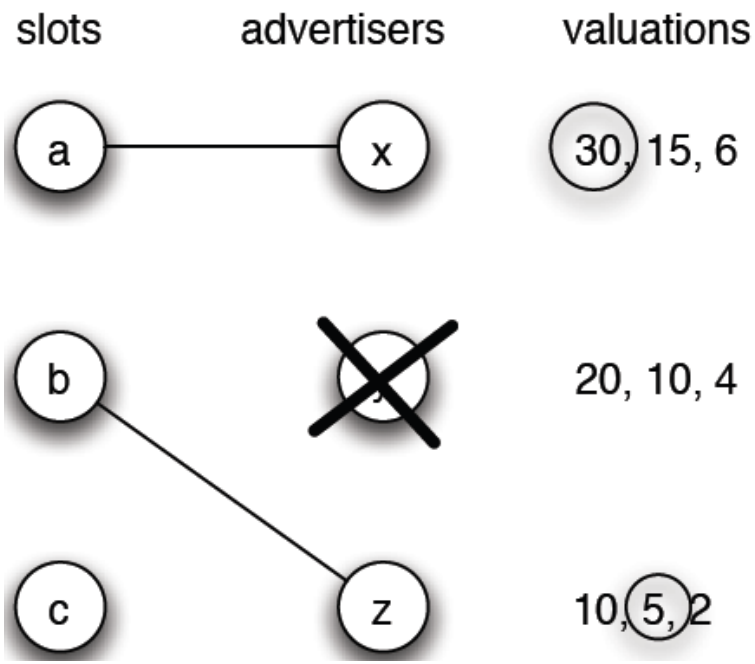
An example

If y weren't there, x would be unaffected, and z would do better by 5-2=3

The total harm y caused is 3

This is the price y should pay

If z weren't there, z cause no harm to x and y



Formulating the "harm"

Let V_B^S be the maximum total valuation over all possible perfect matchings

V_{B-j}^{S-i} be the value of the optimal matching when we take out buyer j and seller i

V_{B-j}^{S} be the value of the optimal matching when we take out buyer j

Let $p_{ij} = V_{B-j}^{S} - V_{B-j}^{S-i}$

This is how everyone do without j (where i is there for sell), and how everyone do with j (where i is taken by j)

The VCG price setting procedure

- Ask buyers to announce valuations for the items (these announcement need not be truthful)
- Choose a socially optimal assignment of items to buyers, i.e., a perfect matching that maximizes the total valuation of each buyer for what they get.
- Charge each buyer the appropriate VCG price: if buyer j receives item i under optimal matching, then charge j a price pij (see pij in the last slide)

Analyze the VCG procedure

- Claim: If items are assigned and prices computed according to the VCG procedure, then truthfully announcing valuations is a dominant strategy for each buyer, and the resulting assignment maximizes the total valuation of any perfect matching of slots and advertisers.

- Two good properties:
  - Maximize total valuation
  - Truthful-telling is a dominant strategy

Proof

- Claim: If items are assigned and prices computed according to the VCG procedure, then truthfully announcing valuations is a dominant strategy for each buyer, and the resulting assignment maximizes the total valuation of any perfect matching of slots and advertisers.

- There are two parts. The second part on maximizing total valuation is easier: as long as they report their valuation truthfully, the total valuation is maximized as the matching is perfect matching

Proof

- For the first part, truthful-telling, we want to show that no one has the incentive to deviate
- That is:
- If buyer j bids true valuation and gets slot i, his payoff is

$$v_{ij} - p_{ij}$$

- Assume he changes his bid and gets slot h, his payoff is

$$v_{hj} - p_{hj}$$

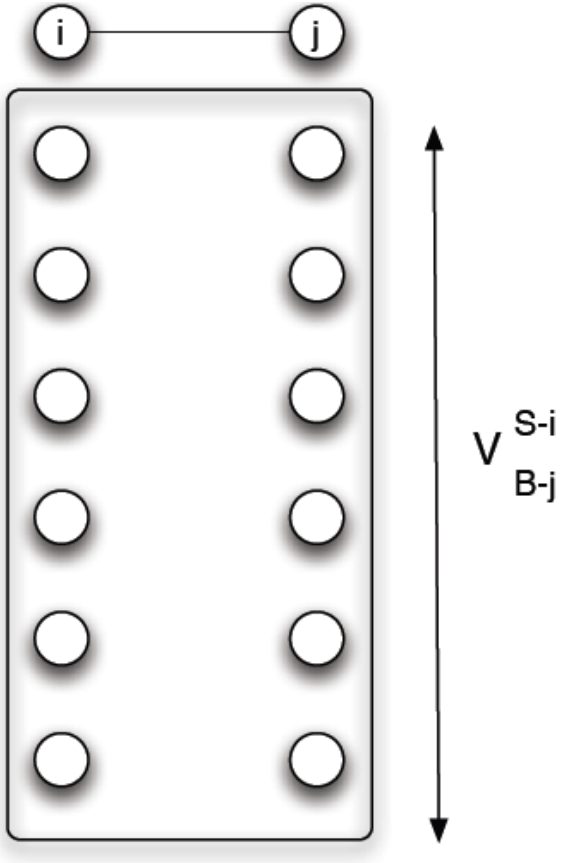
- We need to show

$$v_{ij} - p_{ij} \geq v_{hj} - p_{hj}$$

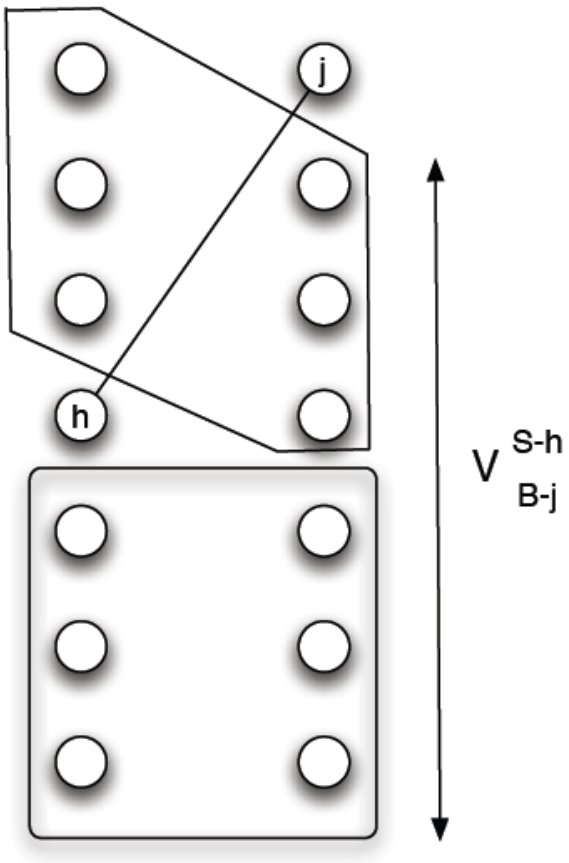
Proof

- We can translate $$v_{ij} - p_{ij} \geq v_{hj} - p_{hj}$$
- to $$v_{ij} - [V_{B-j}^{S} - V_{B-j}^{S-i}] \geq v_{hj} - [V_{B-j}^{S} - V_{B-j}^{S-h}]$$
- which is $$v_{ij} + V_{B-j}^{S-i} \geq v_{hj} + V_{B-j}^{S-h}$$
- We know that $$v_{ij} + V_{B-j}^{S-i} = V_{B}^{S}$$
- and $$v_{hj} + V_{B-j}^{S-h} \leq V_{B}^{S}$$

Proof



(a) $v_{ij} + V_{B-j}^{S-i}$ *is the maximum valuation of any matching.*



(b) $v_{hj} + V_{B-j}^{S-h}$ *is the maximum valuation only over matchings constrained to assign* h *to* j.

Price setting:
Generalized Second-Price Auction (GSP)

The GSP price setting procedure

- Each advertiser j announces a bid consisting of a single number bj, i.e., the price he is willing to pay per click
- After each advertiser submits a bid, GSP awards each slot i to the ith highest bidder, at a price per click equal to the (i+1)st highest bid
- In other words, each advertiser who is shown on the results page is paying a price per click equal to the bid of the advertiser just below them

The GSP price setting procedure

- n slots: click rates r1, r2, …, rn, let's assume in descending order
- n advertisers: bids, b1, b2, …, bn, let's assume in descending order
- Assign r1 to the 1st advertiser, pay by b2, assign r2 to the 2nd advertiser, pay by b3, …

- We can create an artificial advertiser to make this balance
- Set the pay as 0 or a "club joining fee"

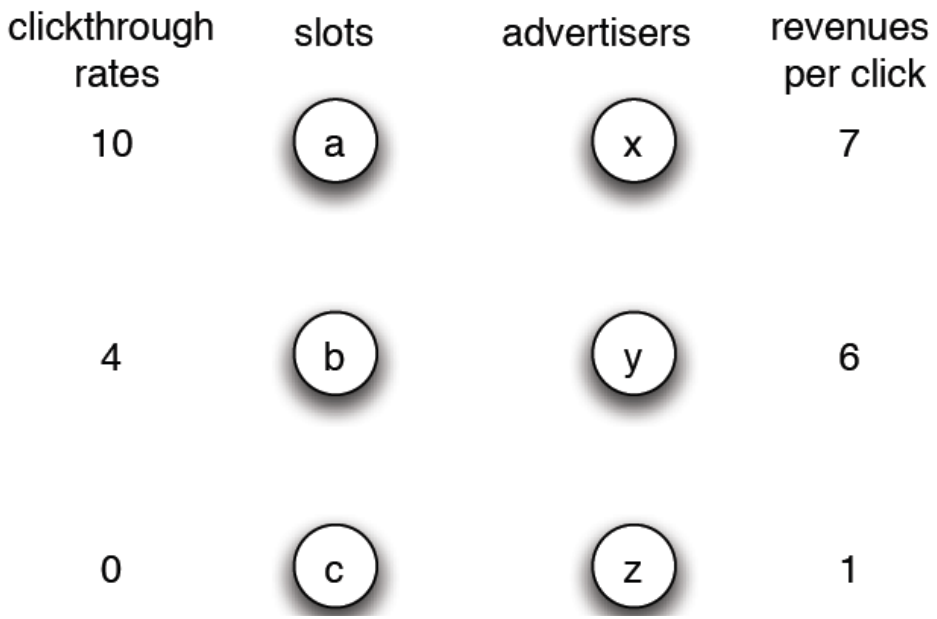
Some basic elements of GSP

- GSP's CPC (cost per click): b_{i+1}
- Payoff = $v_i * r_i - b_{i+1} * r_i$
- v_i is the valuation of advertiser i

Analysis of GSP

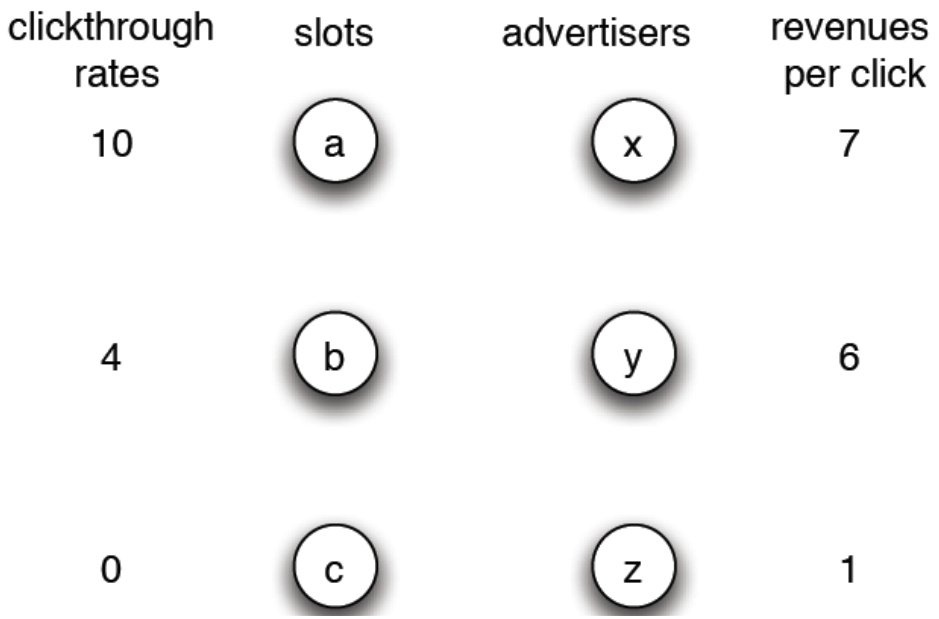
- Truth-telling may not be the dominant strategy for GSP
- In VCG, truth-telling is the dominant strategy

An example for not truth-telling



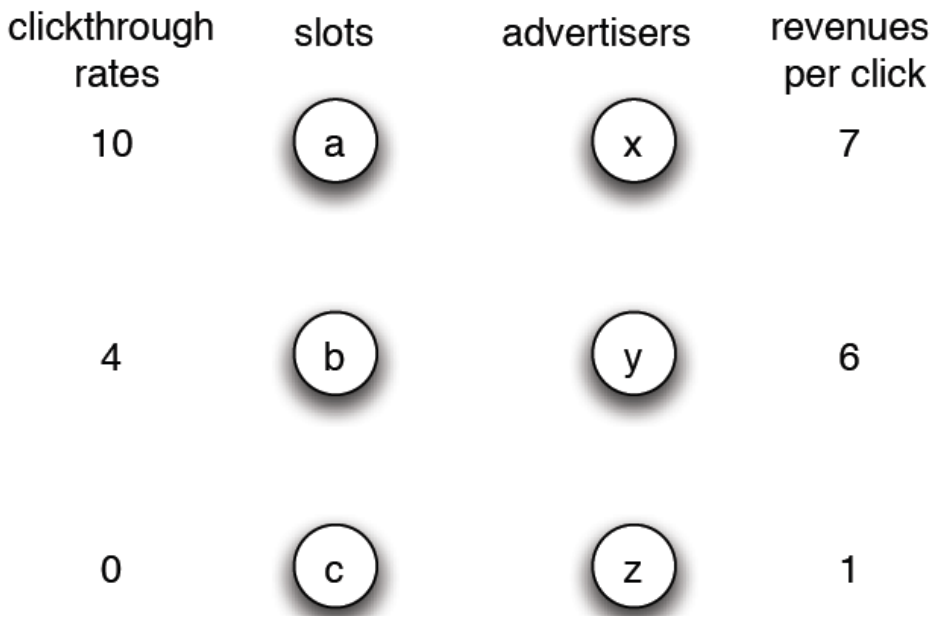
- 3 slots, a, b, c, clickthrough rate 10, 4, 0
- 3 advertisers, x, y, z, every click produces a revenue of 7, 6, 1 for them

An example for not truth-telling



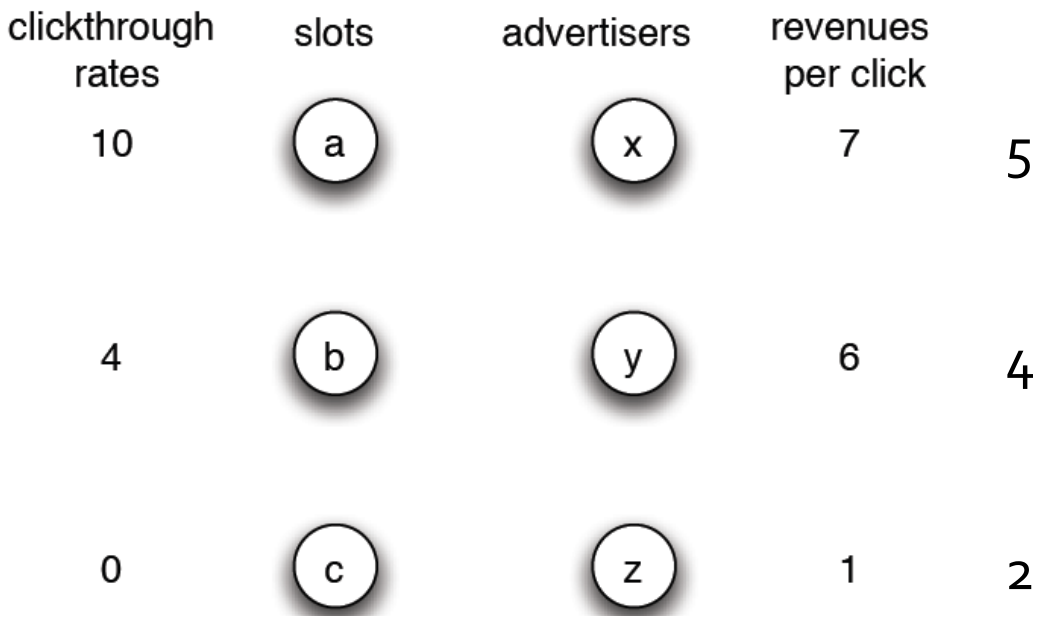
- If truth-telling, x will get slot a, pay at a price of 6, his payoff will be 7*10 - 6*10 = 10
- If x bid 5, what will happen?

An example for not truth-telling



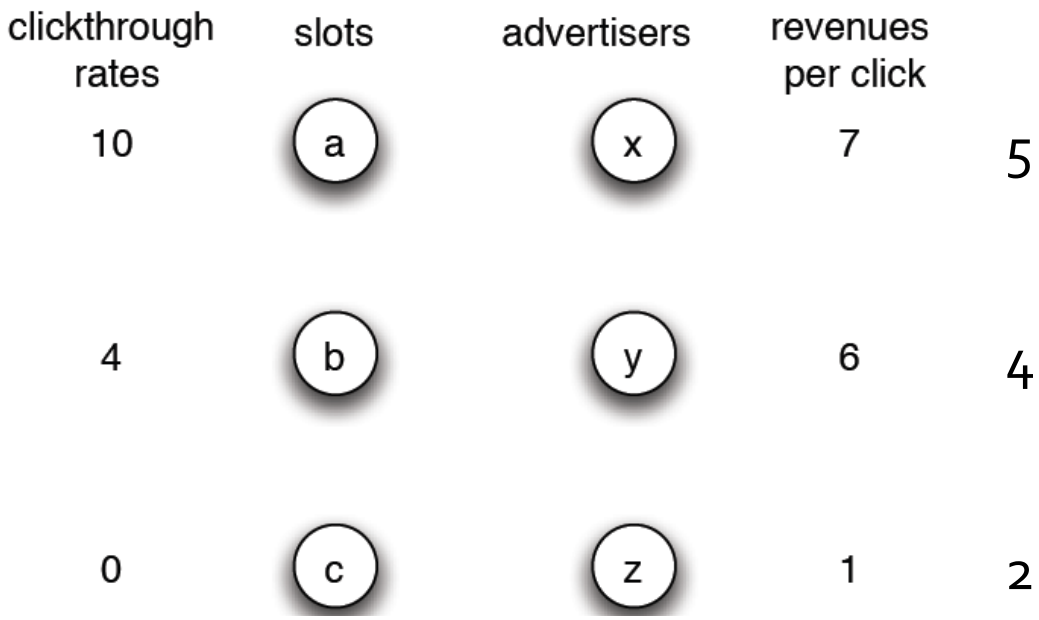
- If truth-telling, x will get slot a, pay at a price of 6, his payoff will be 7*10 - 6*10 = 10
- If x bid 5, what will happen?
- x will get slot b, pay 1, payoff at 7*4-1*4 = 24

Equilibria



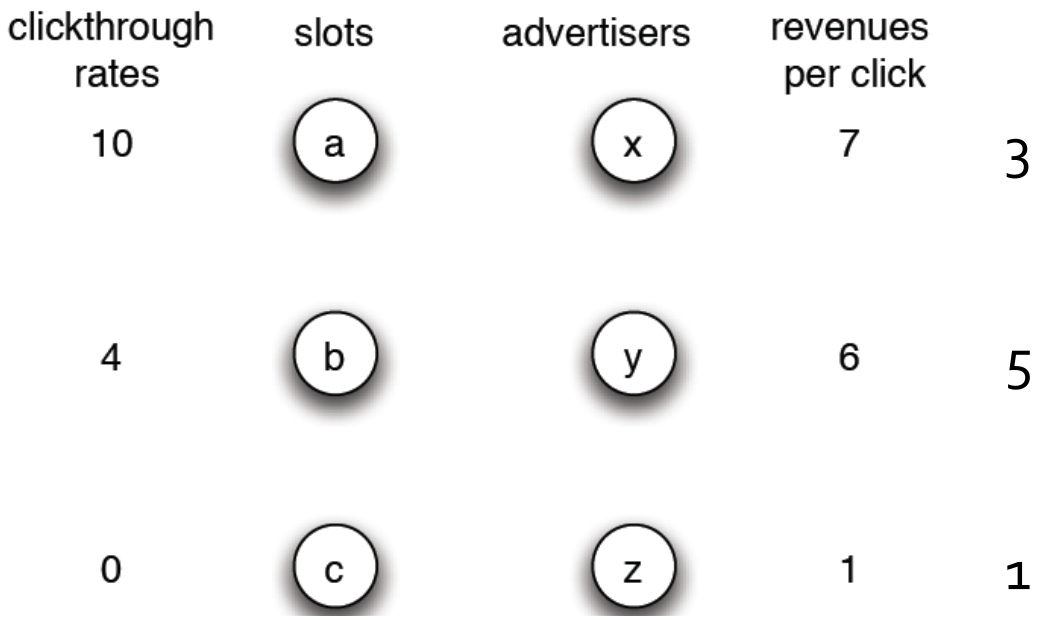
- The bids of 5, 4, 2 form a Nash Equilibrium
- Now, x gets a, with payoff 7*10 – 4*10 = 30. If x bids below 4, he will get b, with payoff 7*4 – 2*4 = 20
- Now, y gets b, with payoff 6*4 – 2*4 = 16. If y bids above 5, he will get a, with payoff 6*10 – 5*10 = 10. If y bids below 2, he will get c, with payoff 0

Equilibria



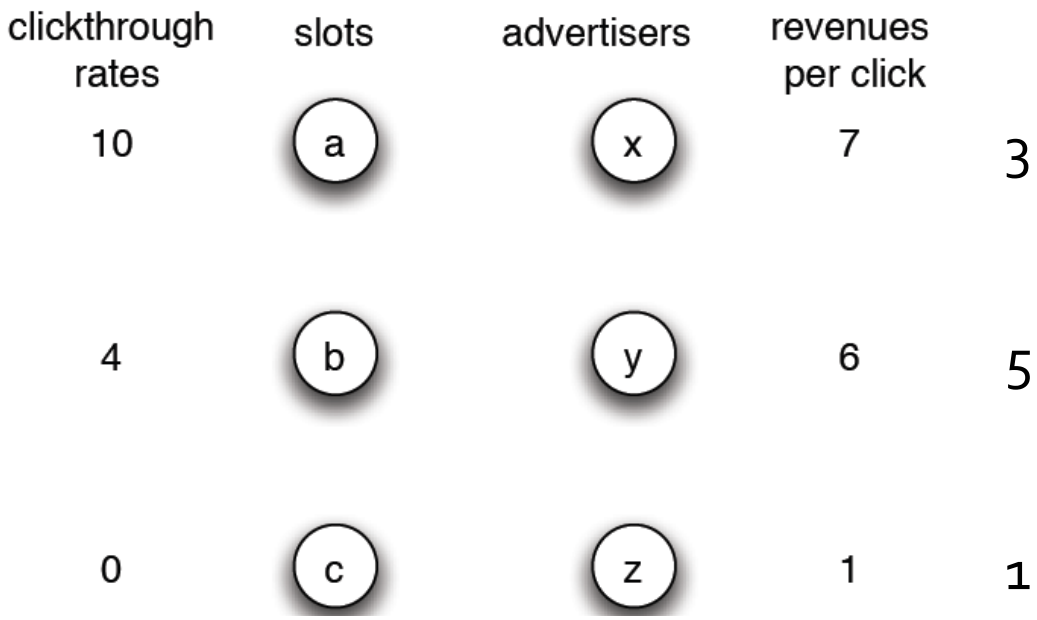
- The bids of 5, 4, 2 form a Nash Equilibrium
- This is also social optimal 70+24+0=94

Equilibria



- The bids of 3, 5, 1 also form a Nash Equilibrium
- Now, x gets b, with payoff 7*4 – 1*4 = 24. If x bids above 5, he will get a, with payoff 7*10 – 5*10 = 20. If x bids below 1, he gets c, with payoff 0
- Now, y gets a, with payoff 6*10 – 3*10 = 30. If y bids below 3, he will get b, with payoff 6*4 – 1*4 = 20.

Equilibria

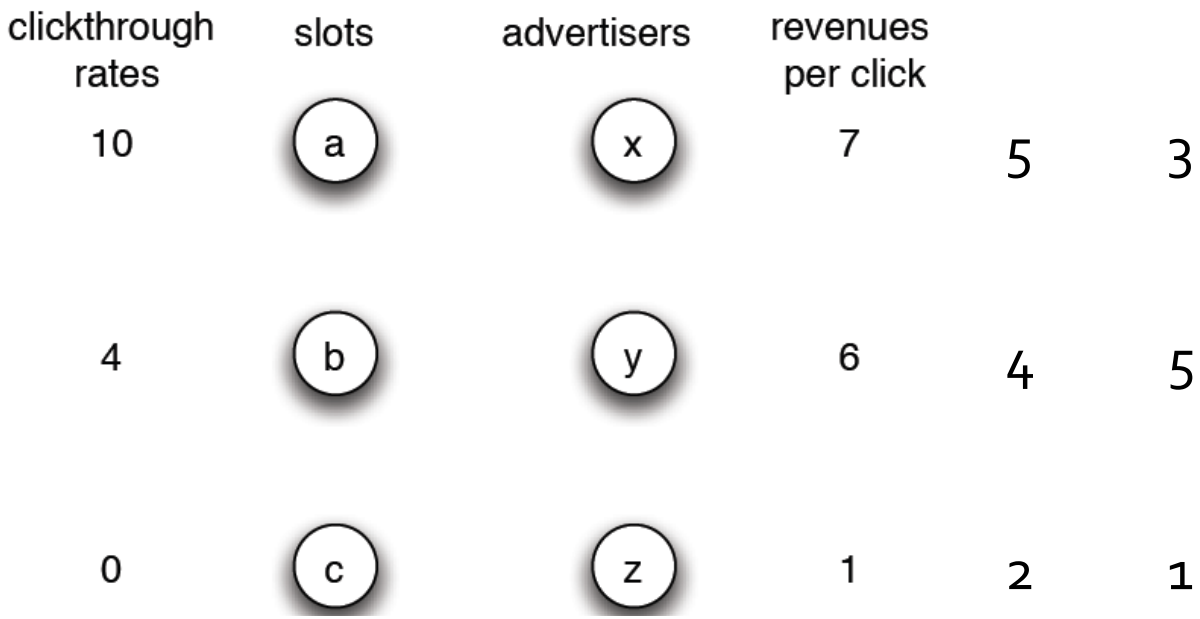


- The bids of 3, 5, 1 also form a Nash Equilibrium
- The social welfare is 7*4 + 6*10 + 1*0 = 88
- This is not social optimal

Some characteristics of GSP

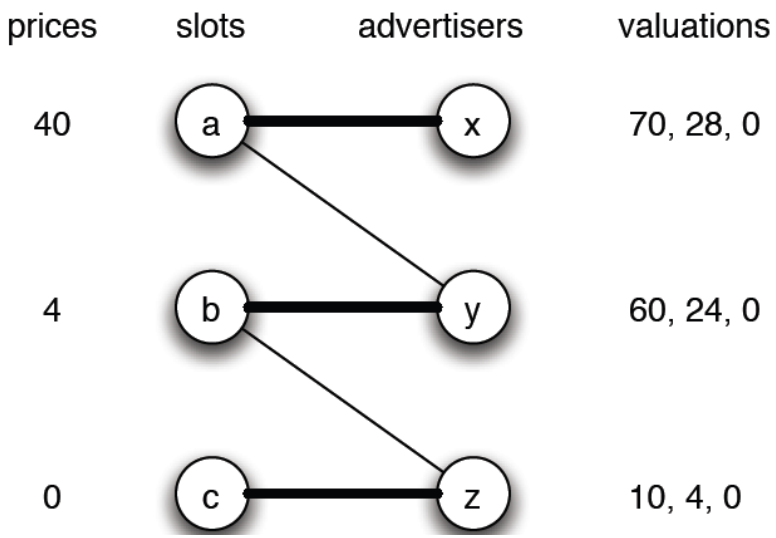
- Truthful-telling is not a dominant strategy
- There can be multiple equilibria, including social optimal and also non-social optimal

The revenue of the search engine of GSP



- For the equilibrium of 5, 4, 2: 4*10 + 2*4 = 48
- For the equilibrium of 3, 5, 1: 3*10 + 1*4 = 34

The revenue of the search engine of VCG



- The harm is (60 – 24) + (4 - 0) = 44
- VCG may not provide a higher revenue to the search engine (of course, this depends on which equilibrium it ends up to in GSP)

Why GSP

- Most search engine currently adopts GSP
- VCG targets on social optimal
- VCG is difficult to compute, difficult to understand by many
- Search engine companies care more on revenue, not social optimal
- VCG has assumptions that, no collusion, the price is only related to clickthrough rate, etc

A summary

- How to auction the advertisement slots
- This is a combination of matching and auction
- VCG has excellent theoretical properties
  - Social optimal
  - Truthful-telling
- GSP (and its variants) has been used widely in practice
  - Easy to understand
  - Search engine may get more