电子科技大学
University of Electronic Science and Technology of China

Cost and Preference in Recommender Systems

Junhua Chen

Data Mining Lab, Big Data Research Center, UESTC
Email：junmshao@uestc.edu.cn
http://staff.uestc.edu.cn/shaojunming

➢ Abstract

In many recommender systems (RS), user's preference is the only factor to be considered in building a RS. Such system take into account the user's preferences part while ignoring another key factor that drives one to purchase – cost/valuation of the item to the user.

➢Outlines

- Preference in RS
- Cost/Valuation
- Modeling Cost & Preference
- Conclusions

➢Outlines

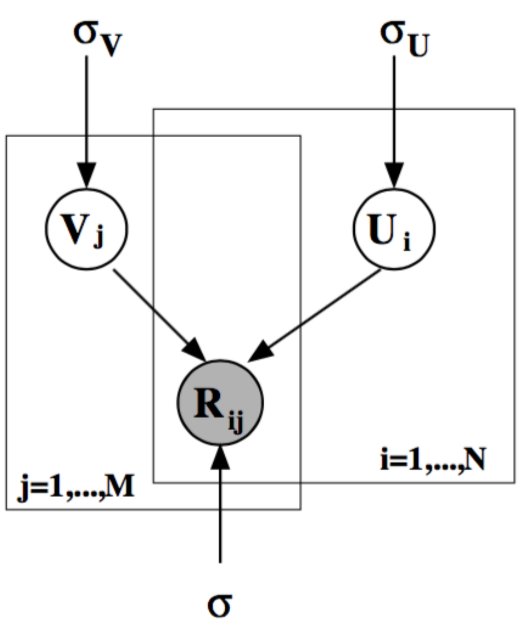
- Preference in RS
- Cost/Valuation
- Modeling Cost & Preference
- Conclusions

Preference

- So many Preference based …
  - Collaborative Filtering (user/item-based)
  - Matrix Factorization
    - Probability Matrix Factorization
    - Tensor Matrix Factorization
    - Time Matrix Factorization
    - Factorization Machine
    - Tag Recommender System
  - Probabilistic Graphical Models
  - Diffusion

Preference

数据挖掘实验室
Data Mining Lab

- Matrix Factorization (my favorite few…)
  - Naïve Matrix Factorization
    - $\mu + b_u + b_i + q_i^T p_u$
  - Time Matrix Factorization
    - $\mu(e) + b_u(e) + b_i(e) + q_i^T(e) p_u(e)$
  - Probability Matrix Factorization
    - 

➢Outlines

- Preference in RS
- Cost/Valuation
- Modeling Cost & Preference
- Conclusions

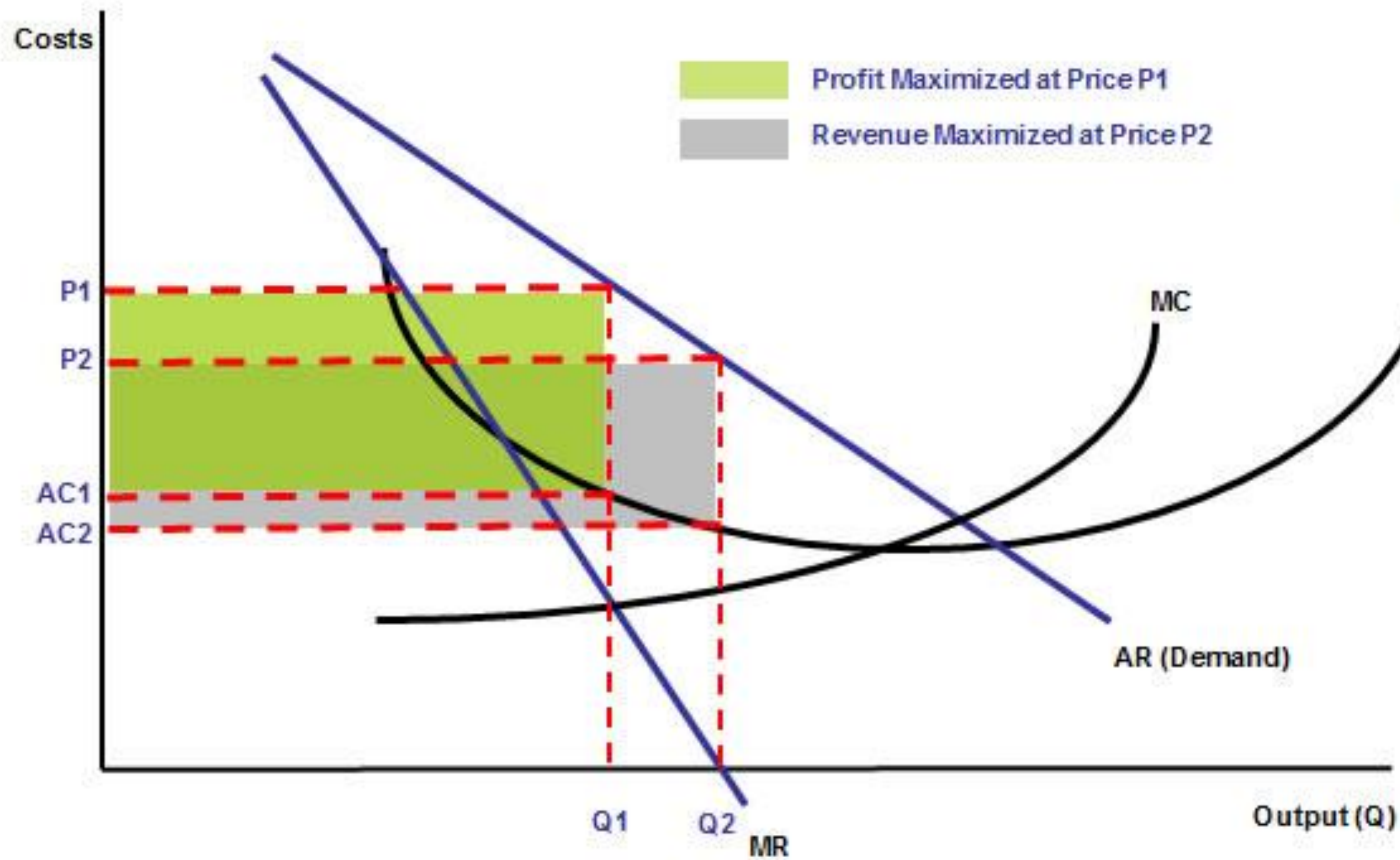
Cost/Valuation

➢Introduction

We talk about Cost/Valuation from the Revenue Maximization (RM) point of view. RM is common in marketing. In the following sections we try to cover the main stream of RM is RS.

Now, let’s talk about money!

➢Profit vs Revenue



Cost/Valuation

数据挖掘实验室
Data Mining Lab

➢Profit vs Revenue - definitions

- R: Revenue = Price * Output
- P: Profit = (Price – Cost) * Output
- TR: Total Revenue
- MR: Marginal Revenue = $\Delta TR/\Delta Q$
- MC: Marginal Cost = $\Delta C/\Delta Q$
- AC: Average total Cost
- AR: Demand
- Green Zone: Profit Maximization
- Grey Zone: Revenue Maximization

数据挖掘实验室
Data Mining Lab

- Cost/Valuation - why

The long-term strategy of any business is to maximize profits because maximizing personal profit is why people start businesses. However, when a small business begins, it may choose to maximize revenue to the detriment of short-term profits so it can build market share and a reputation in the market. Market share is the portion of total sales volume a company controls in a given market.

Cost/Valuation

➢RM in computer-science

- First, how is RM usually appear in computer science.
- Usually, the object of RM is Price(P) * Quantity(Q), P is usually a fixed constant. Q, on the other hand, condition on how consumers pick the item:
  - Indicator {user's value for item i > p of item i}
  - Probability {user's value for item i > p of item}

Cost/Valuation

数据挖掘实验室
Data Mining Lab

➢Finding Seed Group(SG) & Price(P)

- A common RM marketing strategy: The seller selects a set S of buyers and gives them the good for free, and then sets a fixed-price p at which other consumers can buy the item. (denote as ***find SG-P***)
- **Insight**: when a buyer's value for an item is positively influenced by others owing the item.

数据挖掘实验室
Data Mining Lab

➢ Finding Seed Group & Price

- *Mirrokni-WINE 2012*
- Digital good scenario – no quantity limit.
- One's valuation for an item =

its inherent valuation + joint influences from who bought the item.

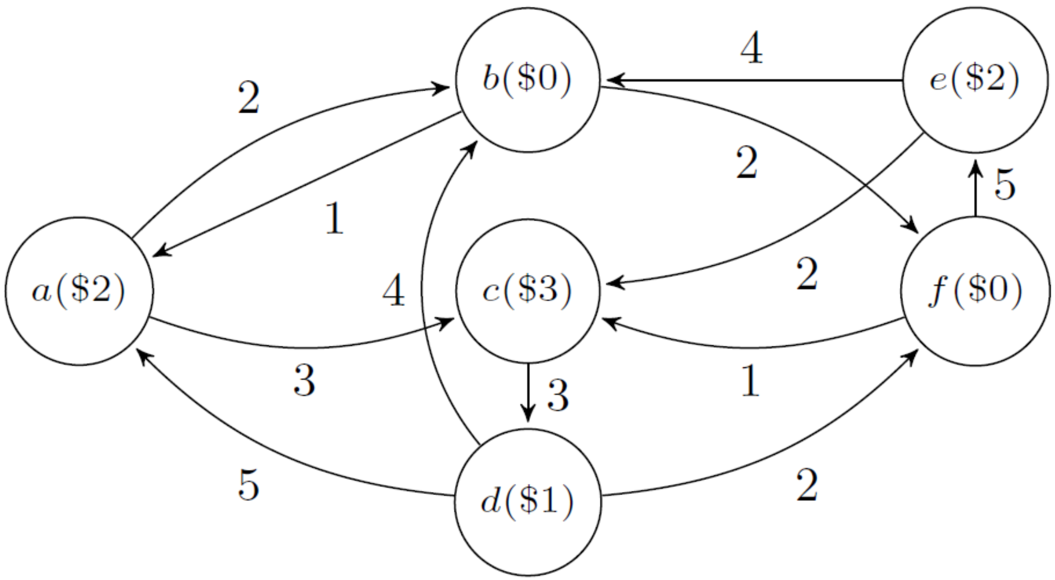
$$v_u = w_u + f_u\left(\sum_{v \in S} w_{uv}\right)$$

- Two factors in diffusion:
  - Pr of infection = Pr ($v_u > p$);
  - Pr of passing the disease = Indicator{user u bought the item}.

Cost/Valuation

➢Finding Seed Group & Price with Quantity

- *Teng-KDD 2015*
- ***Find SG-P*** with Quantity Constraint.
- Interpreting: purchasing real products in Amazon (non digital). Optimizing Display Advertising (in another paper).



➢Finding Seed Group & Price with Quantity

- Problem definition:
  - *Revenue function*:

$$R(u,p,A) = p \times \min\{|\sigma(A)\backslash A|, n - |A|\}$$

  - *RM with a Quantity Constraint*:

Given a monetizing social network **G=(V,X,E,W,F)** , a set of input prices **P**, and a quantity of commodities **n**, the problem is to determine the pricing p_{max}, of the commodity and find a seed group A_{max} as initial customers, where $|A_{max}| < n$, such that the revenue $R(u,p,A)$ is maximized, i.e.,

$$(p_{max}, A_{max}) = argmax_{p,A} R(u,p,A)$$

Cost/Valuation

数据挖掘实验室
Data Mining Lab

➢Finding Seed Group & Price with Quantity

- *Maximum valuation*:

$$X_{max}(v) = \chi_v + F(\sum_{u \in v' sin-neighbors} w_{uv})$$

- *Potential buyers*:

$$X_{max}(v) > p$$

- *Upper bound of maximum revenue*:

$$R_{bound}(n, p) = p \times \min\{n, m_p\}$$

- *Bound of seed Group's Size*:

$$|A| < n - \frac{r_{global}}{p}$$

Cost/Valuation

➢Finding Seed Group & Price with Quantity

- Algorithm **PRUB**

Input: A monetizing social network $G = (V, X, E, W, F)$; a set of input prices P; a quantity of commodities n.
Output: The pricing $p_{\max}$; the seed group $A_{\max}$.

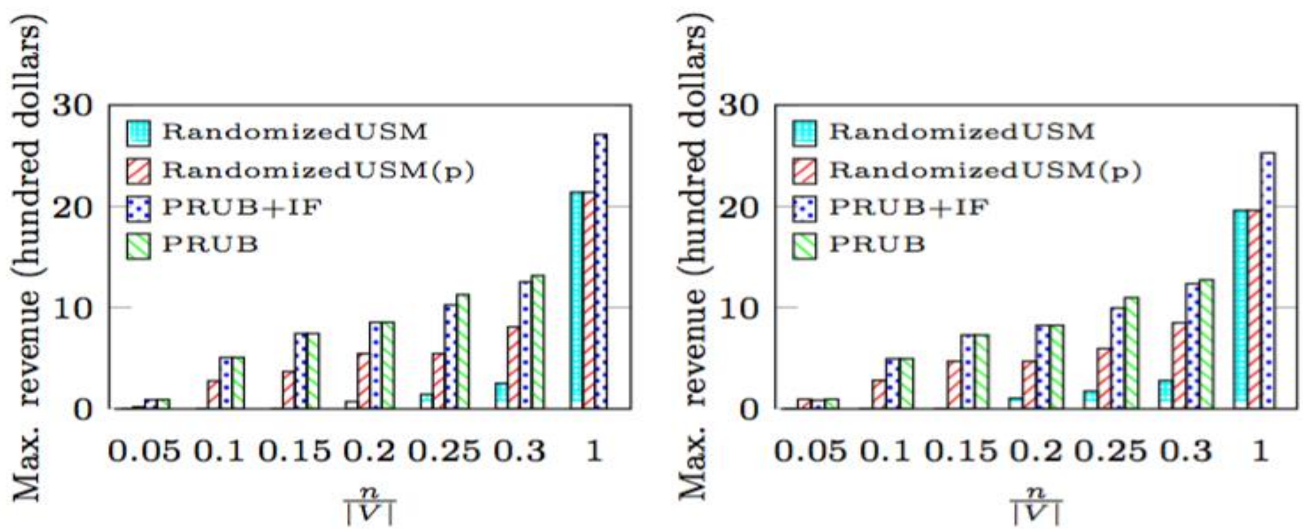
$p_{\max} \leftarrow 0$, $A_{\max} \leftarrow \emptyset$, $r_{\text{global}} \leftarrow 0$
Derive $R_{\text{bound}}(n, p)$ for all $p \in P$
Sort all $p \in P$ descendingly by $R_{\text{bound}}(n, p)$
for $p \in P$ **do**
  if *p is non-candidate pricing* **then**
    return p_{max}, A_{max}
  Enumerate all the seed groups whose size is bounded by Equation (2) (including the size 0)
  Compute $R(n, p, A)$ for those enumerated seed groups
  if *any* $R(n, p, A) > r_{global}$ **then**
    $p_{\max} = p$, $A_{\max} = A$, $r_{\text{global}} = R(n, p, A)$
return p_{max}, A_{max}

- **PRUB-IF**, a more efficient algorithm, greedily select the most important individuals as seeds.

Cost/Valuation

➢Finding Seed Group & Price with Quantity

- Experiments:
  - Datasets: *highschool, dig and facebook* (V & E)
- Results:
  - Revenue results



(a) *highschool*(N) (b) *highschool*(M)

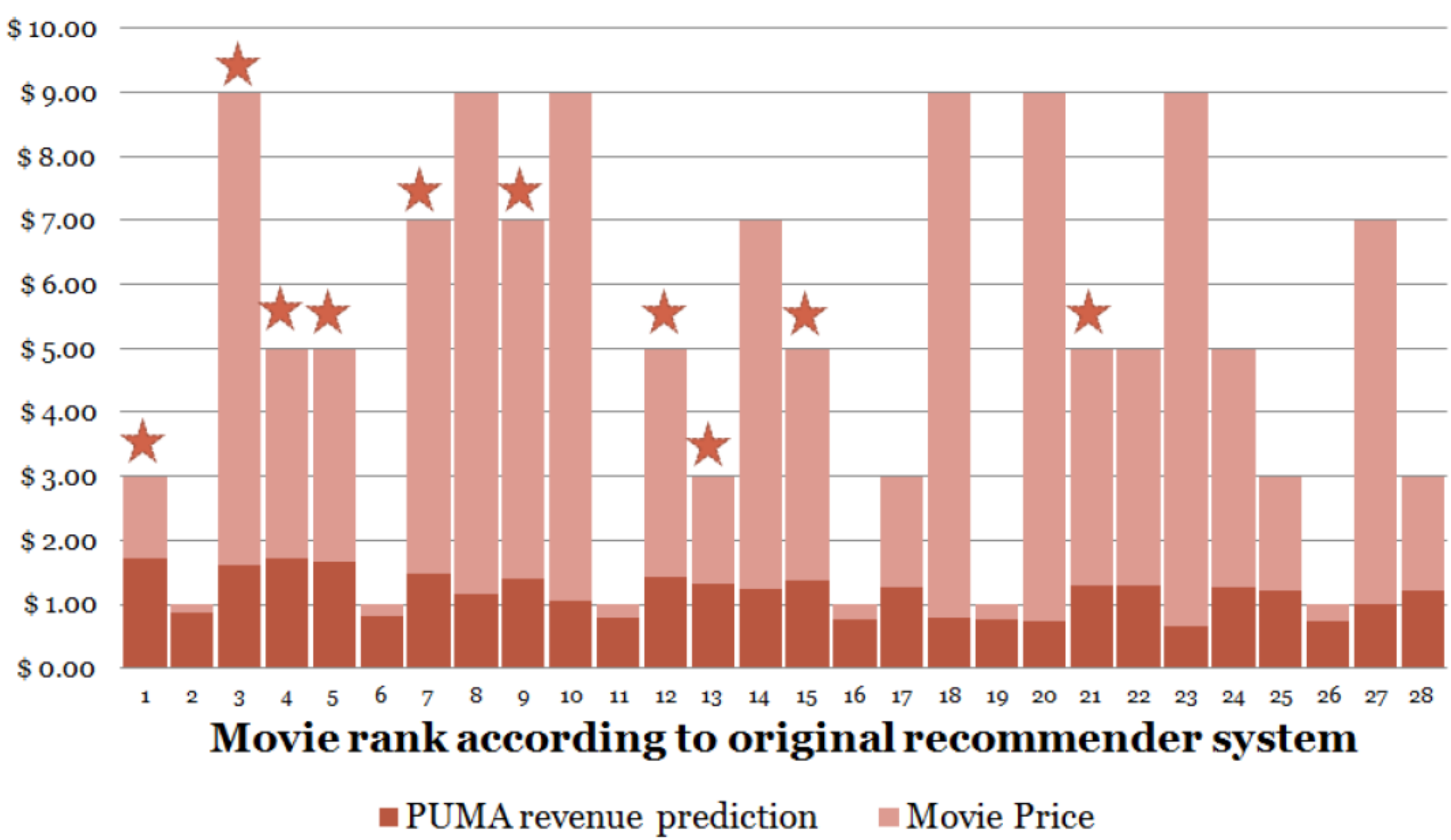
Figure 2: The comparison for maximum revenue.

  - Time comparisons

数据挖掘实验室
Data Mining Lab

➢Other RM Topics

- *Movie Recommender System for Profit Maximization,RecSys13*
- **First look:**



Cost/Valuation

➢Other RM Topics

- *Movie Recommender System for Profit Maximization,RecSys13*
- **Problem**: To increase the revenue based on the recommender system, without any significant drop in user satisfaction.
- **Solution**: *RS* is a *black box*. *Satisfaction* is achieved by survey. The paper model the *probability* of user picking up item m conditioned on its ***rank*** in the RS ranking list and its ***price***.$p(m|r(m), f(m))$.
  - First consider only ranking:
$$p(m|f(m)) = \alpha_1 - \beta_1 \times \ln(f(m))$$
  - Then consider only price:
$$p(m|r(m)) = \alpha_2 - \beta_2 \times \ln(r(m))$$

➢Other RM Topics

- Now the *question* is how to *joint there two probability*, since they are not independent to each other. (simple × dose not make sense, is first Pr is 0.5, second one is 0.5, the final Pr should be 0.5, not 0.5×0.5=0.25, so the paper assume the second Pr, with price, is trained **at its average rank**: n/2+1. By adding a correction term, $\gamma(m)$, we have

$$\alpha_2 + \gamma(m) - \beta_2 \times \ln(r(\frac{n}{2} + 1)) = \alpha_1 - \beta_1 \times \ln(f(m))$$

- by solving the equation and replace it, we have

$$p(m|r(m), f(m)) = \alpha_1 - \beta_2 \times \ln\left(\frac{r(m)}{\frac{n}{2} + 1}\right) - \beta_1 \times \ln(f(m))$$

- Okay, this is all the fun part of this paper, the PM object is simply $(f(m) - c(m)) \times p(m|r(m), f(m))$

Cost/Valuation

数据挖掘实验室
Data Mining Lab

➢Other RM Topics

- *Optimizing Display Advertising in Online Social Networks, W315*
- **Problem**: Social Display Optimization – given a tuple (U,B,p) set of users U, bound B, influence function p (per user), the goal is to find a strategy for showing ad to B users so as to maximize the expected clicks.

- **Solution**:
  Display strategy
  - Linear influence:

$$p_u(S) = c_u + \sum_{v \in S} w(u, v)$$

Cost/Valuation

➢Other RM Topics

- Independent Cascade Model:

$$p_u(s) = 1 - \prod\nolimits_{s} (1 - p(u, v))$$

where P(u,v) is the influences of pair of users u,v.

- Concave influence:

$$p_u(S) = g(\sum\nolimits_{v \in S} w(u, v))$$

where $g(x) = x^d\ or\ logx.$

- Deterministic threshold function:

$$p_u(S) = \begin{cases} 1, \sum\nolimits_{v \in S} w(u, v) > T_u \\ 0, otherwise \end{cases}$$

➢Other RM Topics

- **Algorithm**
  - *Largest probability greedy*: greedily select the $p_u(S)$
  - *Most influential greedy*: greedily select the $infl(u) = \sum_{u \in U} p_v(\{u\})$
  - *Adaptive Hybrid Heuristic*: greedily select the $infl(u) \times p_u(S)$

Adaptive Hybrid Heuristic

Input: (U, B, p_u).

Output: A sequence $(a_1, a_2, \ldots, a_B)$, and set S.

Goal: Maximize $E[|S|]$.

1. **Initialize:** $S = \emptyset, A = \emptyset$
2. **For** $i := 1$ **to** B **do**
3. **Let** $a_i \in U \setminus A$ be the user maximizing $p_{a_i}(S) \times$ top-infl(a_i)
4. **Let** $A := A \cup \{a_i\}$.
5. With probability $p_{a_i}(S)$, **let** $S = S \cup \{a_i\}$.

Figure 2: Adaptive Hybrid Heuristic.

➢Other RM Topics

- *Two-stage Heuristic*: first a step use max infl, the rest step use max probability.

Two-stage Heuristic

Input: (U, B, p_u).

Output: A sequence $(a_1, a_2, \ldots, a_B)$, and set S.

Goal: Maximize $E[|S|]$.

1. **Initialize:** $S = \emptyset$, $A = \emptyset$
2. **For** $i := 1$ **to** αB **do**
3. **Let** $a_i \in U \backslash A$ be the user maximizing top-infl(u)
4. **Let** $A := A \cup \{a_i\}$.
5. With probability $p_{a_i}(S)$, **let** $S = S \cup \{a_i\}$.
6. **For** $i := \alpha B + 1$ **to** B do **do**
7. **Let** $a_i \in U \backslash A$ be the user maximizing $p_{a_i}(S)$
8. **Let** $A := A \cup \{a_i\}$.
9. With probability $p_{a_i}(S)$, **let** $S = S \cup \{a_i\}$.

➢Other RM Topics

- *Show Me the Money: Dynamic Recommendations for Revenue Maximization, VLDB14*
- ***Competition*** and ***Saturation Effects***
  - *Competition*:

  A user u adopts an item i at time t is conditioned on u **not** adopting anything from i's class **before**, and once u adopts I, she will **not** adopt anything from that class **again** in the horizon.

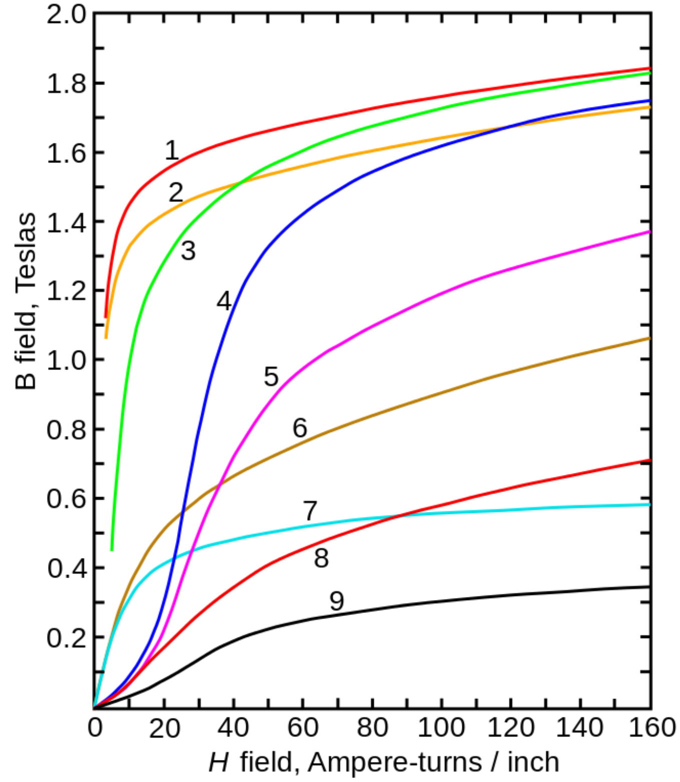
  - *Saturation effects*:

  Repeated suggestion may lead to a **boost** in adoption probability, but repeating too frequently may **backfire**, as people have a tendency to develop **boredom**.

➢Other RM Topics

- More *saturation effects*

Seen in some magnetic materials, saturation is the state reached when an increase in applied external magnetic field H cannot increase the magnetization of the material further, so the total magnetic flux density B more or less levels off.



➢Other RM Topics

- The **memory** of user u on item i at any time t with a RS strategy S is:

$$M_S(u,i,t) := \sum_{j\in C(i)} \sum_{\tau=1}^{t-1} \frac{X_S(u,j,\tau)}{t-\tau}$$

where $X_S(u,j,\tau)$ is an indicator variable taking on 1 if $(u,j,\tau) \in S$ and 0 otherwise.

- The saturation effects can be modeled as:

$$q(u,i,t) \times \beta_i^{M_S(u,i,t)}$$

where β is the saturation factor, small β means greater saturation effect. $q(u,i,t)$ is (primitive) adoption Pr.

Other RM Topics

- Overall, the picking up probability would be

$$q_S(u,i,t)$$
$$= q(u,i,t) \times \beta_i^{M_S(u,i,t)} \times \prod_{(u,j,t)\in S: j\neq i, C(i)=C(j)} (1-q(u,j,t))$$
$$\times \prod_{(u,j,\tau)\in S: \tau<t, C(j)=C(i)} (1-q(u,j,\tau))$$

Where the first part is saturation effect and the later two part is competition.

➢Outlines

- Preference in RS
- Cost/Valuation
- Modeling Cost & Preference
- Conclusions

Cost & Preference

数据挖掘实验室
Data Mining Lab

- How to infer the valuation function for user.
  - We assume that what drives one to pick something is conditioned on one's preference and its valuation for the object.
  - We hope to face this challenge with helps from inferring network.

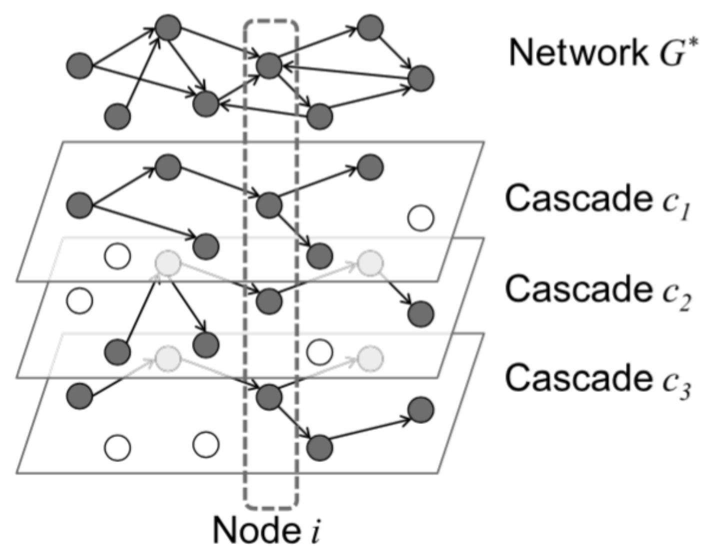
Cost & Preference

➤Inferring Network

- *Inferring Networks of Diffusion and Influence, TKDD 2012*
- Cascade: in data, it is series, $(v, t_v)_C$.
- Hypothesis: a cascade is a Tree.
- Diffusion:

$$p(c|T) = \prod_{(u,v)\in E_T} \beta\, P_c(u,v) \prod_{u\in V_T,(u,x)\in E\setminus E_T} (1-\beta)$$

Notice $|E_T| = |V_T| - 1$, essentially,
$p(c|T)$ is independent to β.



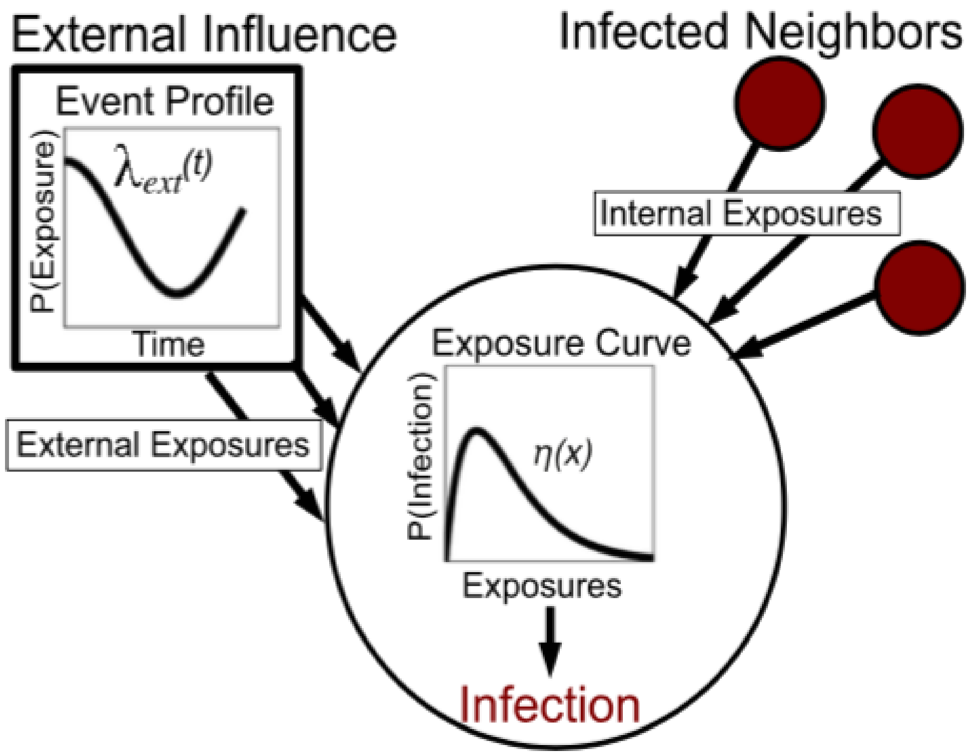
➢ Inferring Network

- $\boldsymbol{\varepsilon - edge}$, external Influence. To account for such phenomena when a cascade jumps across the network, the author creates an additional node x that represents an external influence and can infect any other node u with small probability ε. (external influence can be interpreted as mainstream media).

$$P'_c(u,v) = \begin{cases} \beta P_c(t_v - t_u) & \text{if } t_u < t_v \text{ and } (u,v) \in E_T \cap E & (u,v) \textit{ is network edge} \\ \varepsilon P_c(t_v - t_u) & \text{if } t_u < t_v \text{ and } (u,v) \in E_T \cap E_\varepsilon & (u,v) \textit{ is } \varepsilon\textit{-edge} \\ 1-\beta & \text{if } t_v = \infty \text{ and } (u,v) \in E \backslash E_T & v \textit{ is not infected, network edge} \\ 1-\varepsilon & \text{if } t_v = \infty \text{ and } (u,v) \in E_\varepsilon \backslash E_T & v \textit{ is not infected, } \varepsilon\textit{-edge} \\ 0 & \text{else } (i.e., t_u \geq t_v). \end{cases}$$

➢Network Diffusion

- *Information Diffusion and External Influence in Networks,KDD12*
- There are two different ways of how information reaches a person in a network, through **connections in our social network** and **external out-of-network sources**.



consider a twitter scenario, one can get information from its friends, or learn the information from external big media like CNN or ABC.

➢Network Diffusion

- How? To model *internal influence* and *external influence*.
- From the figure, *Pr of infection* is a function of exposures, which means, the author transform all influence / information (internal or external) into **exposure** concept.

❖ **Modeling** the internal exposures:

- **Hazard function**: they describe a distribution of the length of time it takes for an event to occur.
- Let

$$\lambda_{int}(t)dt \equiv P(i\ exposes\ j\ \in [t, t+dt] | i\ hasn't\ exposed\ j\ yet)$$

λ_{int} effectively models how long it takes a node to notice one of its neighbors becoming infected.

➢Network Diffusion

- The **expected number** of internal **exposures** a node i has received by time t, is defined by the sum of the cumulative distribution functions of exposures.

$$\Lambda_{int}^{(i)}(t) = \sum_{j:j\ is\ i's\ inf.neighbor} P(j\ exposed\ i\ before\ t)$$

$$= \sum_{j} [1 - \exp(-\int_{\tau_j}^{t} \lambda_{int}(s - \tau_j) ds)]$$

Where hazard function use exponential failure distribution.

- *PS: more hazard function,*

$$\lambda(t) = \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)}$$

R(t): the Pr of no failure before time t.

Network Diffusion

Modeling the external exposures:

- The external source **cannot be observed**. The source varies in intensity over time, the paper define it as event profile:

$$\lambda_{ext}(t)dt \equiv P(i\ receives\ exposure\ \in [t, t+dt])$$

- all nodes have the **same Pr of receiving** an external exposure for any point in time. The paper model the arrival of exposure as **a binomial distribution**.

$$P_{exp}(n;t) = \binom{\frac{t}{dt}}{n} (\lambda_{ext}dt)^n \times (1-\lambda_{ext}dt)^{\frac{t}{dt}-n}$$

➢Network Diffusion

- to relax the constraint that λ_{ext} is constant, using the average of it.

$$P_{exp}(n;t) \approx \binom{\frac{t}{dt}}{n} (\frac{\Lambda_{ext}(t)}{t} dt)^n \times (1 - \frac{\Lambda_{ext}(t)}{t} dt)^{\frac{t}{dt}-n}$$

- Finally, users are receiving **both external and internal** exposures at the same time, do the **average**.

$$P_{exp}^{(i)}(n;t) \approx \binom{\frac{t}{dt}}{n} (\frac{\Lambda_{ext}(t) + \Lambda_{int}^{(i)}(t)}{t} dt)^n \times (1 - \frac{\Lambda_{ext}(t) + \Lambda_{int}^{(i)}(t)}{t} dt)^{\frac{t}{dt}-n}$$

➢Network Diffusion

❖ **Modeling** the exposure curve

$\eta(x) \equiv P(node\ i\ is\ infected\ immediately\ after\ x^{th}\ exposure)$

$$\eta(x) = \frac{\rho_1}{\rho_2} \cdot x \cdot \exp\left(1 - \frac{x}{\rho_2}\right)$$

- Finally, the **cumulative distribution function** of the infection probability:

$$F^{(i)}(t) = \sum_{n=1}^{\infty} P[i\ has\ n\ exp.] \times P[i\ inf.|i\ has\ n\ exp.]$$

$$= \sum P_{exp}^{(i)}(n;t) \times [1 - \prod_{k=1}^{n} [1 - \eta(k)]]$$

➢Outlines

- Preference in RS
- Cost/Valuation
- Modeling Cost & Preference
- Conclusions

Conclusions

- Conclusions and future works

  Cost and Preference in Recommender system. It aims at understanding what drive a consumer to purchase. We hope to learn through users' history purchased data to learn ones' preference, through social network to learn users' valuation (or cost) of purchasing. Finally, inferring a valuation function on time that determined the probability of one picking up an item under the influence of cost and preference.

Conclusions

- Acknowledgment

Conclusions

数据挖掘实验室
Data Mining Lab

- Questions

Thanks

Junhua Chen
Data Mining Lab, Big Data Research Center
Junhuachen.jhc@gmail.com