Group-based ranking method for online rating systems with spamming attacks

JIAN GAO

gaojian08@hotmail.com

CompleX Lab, Web Sciences Center, UESTC

March 28, 2015

Collaborators

Yuwei Dong, Shimin Cai

Mingsheng Shang, Tao Zhou

Web Sciences Center, University of Electronic Science and Technology of China,

Chengdu 611731, P.R.C.

Outline

- Backgrounds
- Method & Data
- Main results
- Conclusions

Backgrounds

Backgrounds

Which to buy?

Backgrounds

Reputation evaluation system

店铺动态评分：（所属行业：食品/保健）

商品与描述相符：	4.8分	比同行业平均水平	高12.47%
商家的服务态度：	4.8分	比同行业平均水平	高6.17%
商家发货的速度：	4.8分	比同行业平均水平	高18.56%

4.8分 ★★★★★共35873人

5 分 89.77% (32203人)
4 分 7.66% (2748人)
3 分 1.37% (492人)
2 分 0.44% (158人)
1 分 0.76% (272人)

Tmall.com 淘宝商城

店铺动态评分：（所属行业：食品/保健）

商品与描述相符：	4.8分	与同行业平均水平	持平
商家的服务态度：	4.8分	比同行业平均水平	低0.69%
商家发货的速度：	4.8分	比同行业平均水平	低0.51%

4.8分 ★★★★★共98769人

5 分 88.23% (87145人)
4 分 8.69% (8579人)
3 分 1.81% (1783人)
2 分 0.36% (358人)
1 分 0.92% (904人)

ebay

Customer Reviews

(2,251)
3.6 out of 5 stars

5 star	986
4 star	399
3 star	220
2 star	199
1 star	447

See all 2,251 customer reviews

"Looks good, works great, very easy to install."
snowmom | 261 reviewers made a similar statement

"I love the matt finish, no glare, no fingerprints, no smudges."
rJacquieM | 140 reviewers made a similar statement

"If you want a good screen protector for your phone that doesn't have the anti glare/matte finish then you'd be okay purchasing this."
Shan Shan | 107 reviewers made a similar statement

amazon.com

Backgrounds

Spamming attacks

China's Internet "Water Army" (Shuijun) - Opinion Spammers

- You can hire people to write and post fake reviews or comments, and even bribe staff at review, forum
- 'Water Army' Whistleblower Threatened, January 7, 2011, People's Daily.
- The Chinese Online "Water Army", June 25, 2010, Wired.com.
- If you read Chinese, see this description from Baidu Baike at baidu.com.

Backgrounds

How to deal with

the spamming problem

in online rating systems?

Jindal N. and Liu B., ICWWW, 2007, pp. 1189-1190.

Pan W., Xiang E.-W., Liu N.-N. and Yang Q., AAAI, 2013, pp. 39-55.

Backgrounds

Building users' reputation systems

Online rating systems

	o1	o2	o3	o4	o5
u1	4	5	3	4	-
u2	-	4	4	2	5
u3	3	4	-	5	3
u4	5	-	-	4	3
u5	3	4	5	-	3

Backgrounds

Building users' reputation systems

- Iterative refinement (IR) method

  Laureti P., Moret L., Zhang Y.-C. and Yu Y.-K., EPL, 75 (2006) 1006.

- An improved IR method

  de Kerchove C. and Van Dooren P., arXiv:0711.3964 (2007)

- The correlation-based ranking (CR) method

epl A Letters Journal Exploring the Frontiers of Physics

May 2011

EPL, **94** (2011) 48002
doi: 10.1209/0295-5075/94/48002

www.epljournal.org

A robust ranking algorithm to spamming

YAN-BO ZHOU[1], TING LEI[1] and TAO ZHOU[2(a)]

[1] *Département de Physique, Université of Fribourg - CH-1700 Fribourg, Switzerland*
[2] *Web Sciences Center, University of Electronic Science and Technology of China - 610054 Chengdu, PRC*

EPL, best of 2011 collection

Backgrounds

The correlation-based ranking (CR) method

User rating: $r_{i\alpha}$ Object quality: Q_{O_α} Reputation: Cu_i

$$Qo_\alpha = \frac{\sum_{i \in U_\alpha} Cu_i r_{i\alpha}}{\sum_{i \in U_\alpha} Cu_i}. \quad (1)$$

$$Corr_i = \frac{1}{ku_i} \sum_{\alpha \in O_i} \left(\frac{r_{i\alpha} - \overline{r_i}}{\sigma_{r_i}} \right) \left(\frac{Qo_\alpha - \overline{Qo_i}}{\sigma_{Qo_i}} \right), \quad (2)$$

$$Cu_i = \begin{cases} corr_i, & \text{if } corr_i \geqslant 0, \\ 0, & \text{if } corr_i < 0. \end{cases} \quad (3)$$

$$|Qo - Qo'| = 1/|O| \sum_{\alpha \in O} (Qo_\alpha - Qo'_\alpha)^2, \quad (4)$$

Backgrounds

- Iterative algorithm with reputation redistribution (IARR)

OPEN ACCESS Freely available online

PLOS ONE

Ranking Reputation and Quality in Online Rating Systems

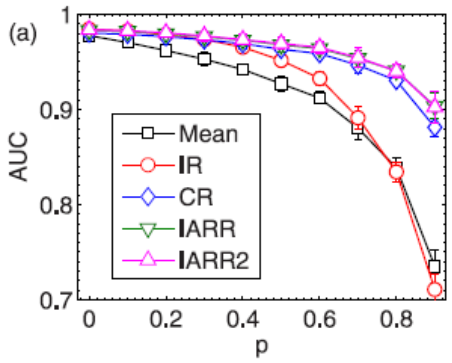
Hao Liao[1], An Zeng[1]*, Rui Xiao[1], Zhuo-Ming Ren[1,2], Duan-Bing Chen[1,3], Yi-Cheng Zhang[1]

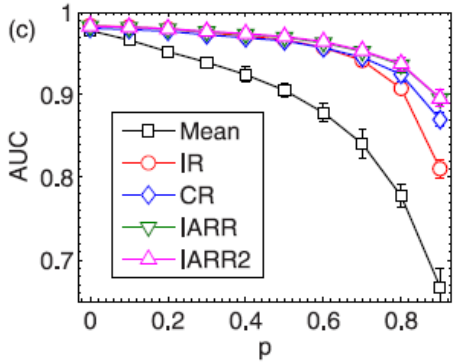
1 Department of Physics, University of Fribourg, Fribourg, Switzerland, 2 Research Center of Complex Systems Science, University of Shanghai for Science and Technology, Shanghai, China, 3 Web Sciences Center, University of Electronic Science and Technology of China, Chengdu, China

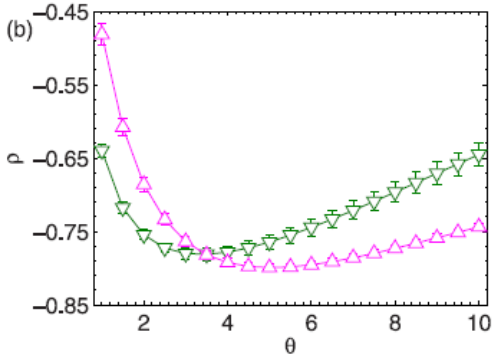
$$Q_\alpha = \frac{\sum_{i \in U_\alpha} R_i r_{i\alpha}}{\sum_{i \in U_\alpha} R_i} \quad (1)$$

$$TR_i = \frac{1}{k_i} \sum_{\alpha \in O_i} \left(\frac{r_{i\alpha} - \bar{r}_i}{\sigma_{r_i}}\right)\left(\frac{Q_\alpha - \bar{Q}_i}{\sigma_{Q_i}}\right) \quad (2)$$

$$R_i = TR_i^\theta \frac{\sum_j TR_j}{\sum_j TR_j^\theta} \quad (3)$$





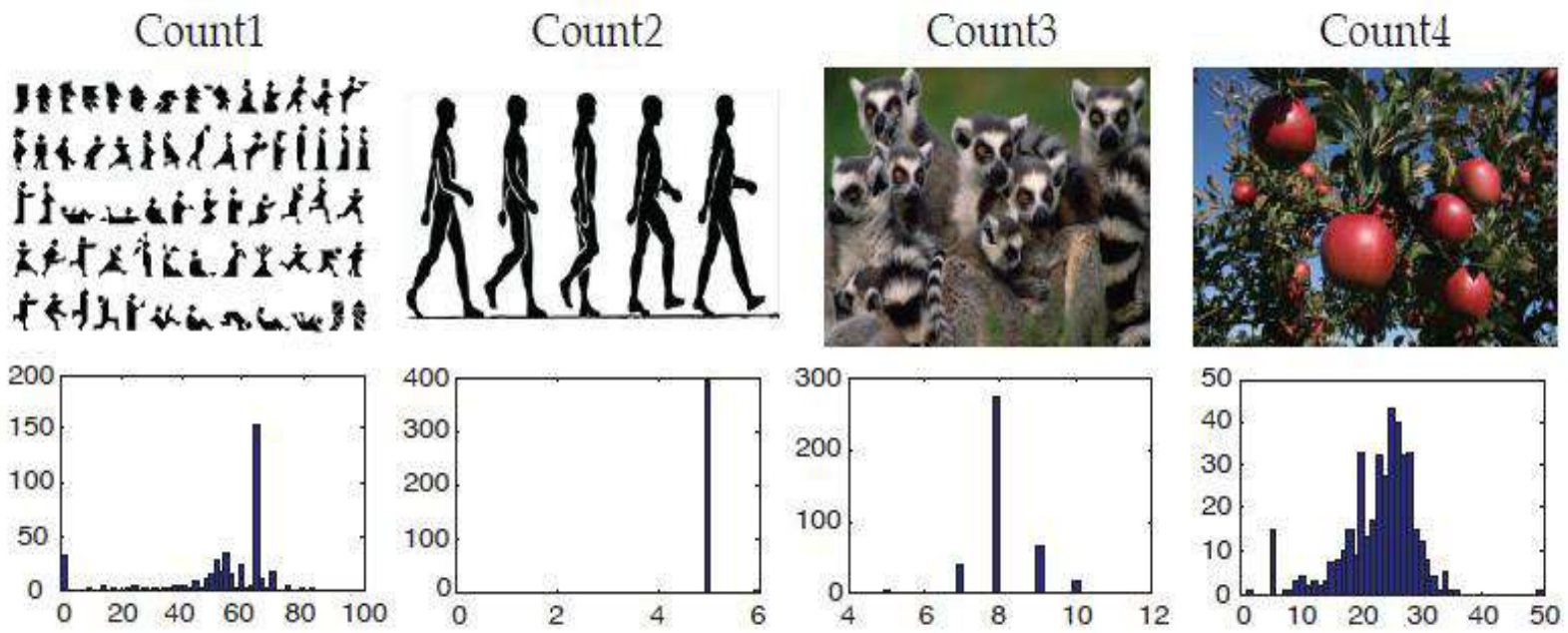


Backgrounds

- Previous works
  - Estimate quality of objects
  - Estimate reputation of users
- Key assumption
  - Every object is associated with a most objective rating that best reflect its quality
- Advantages
  - Effective with artificial data
  - Robust spamming attacks

Backgrounds

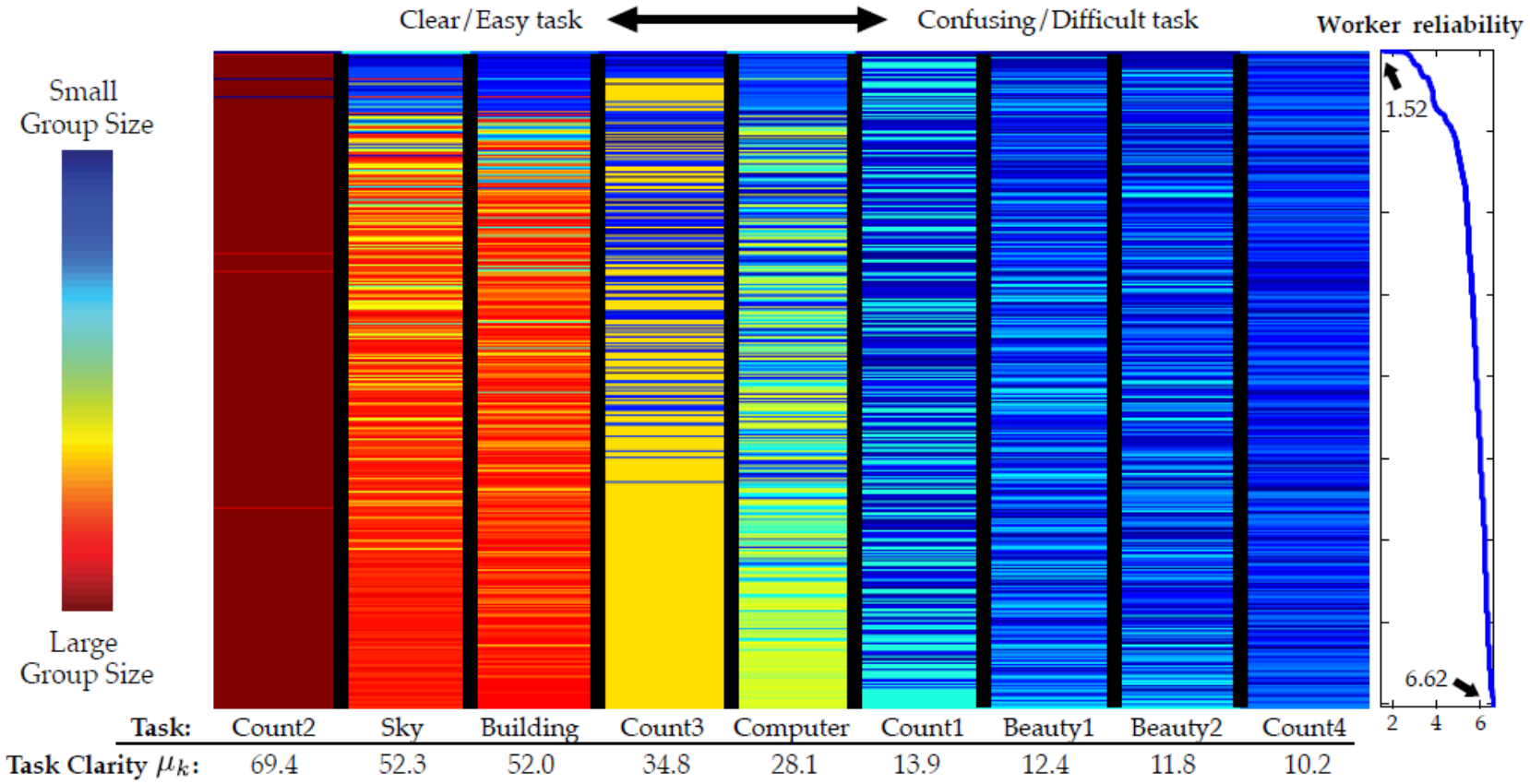
- The group structure of schools of thought
  - Identifying reliable workers and unambiguous tasks in data collection



Tian Y, Zhu J. ACM SIGKDD. 2012. 226-234.

Backgrounds

- The group structure of schools of thought



Backgrounds

- What can we learn ?
  - The more clear of a task, the larger the group size is
  - The larger the group size is, the higher the reliability is

a random rating to objects with confusing quality should be acceptable

a biased rating to objects with clear quality should be unreasonable

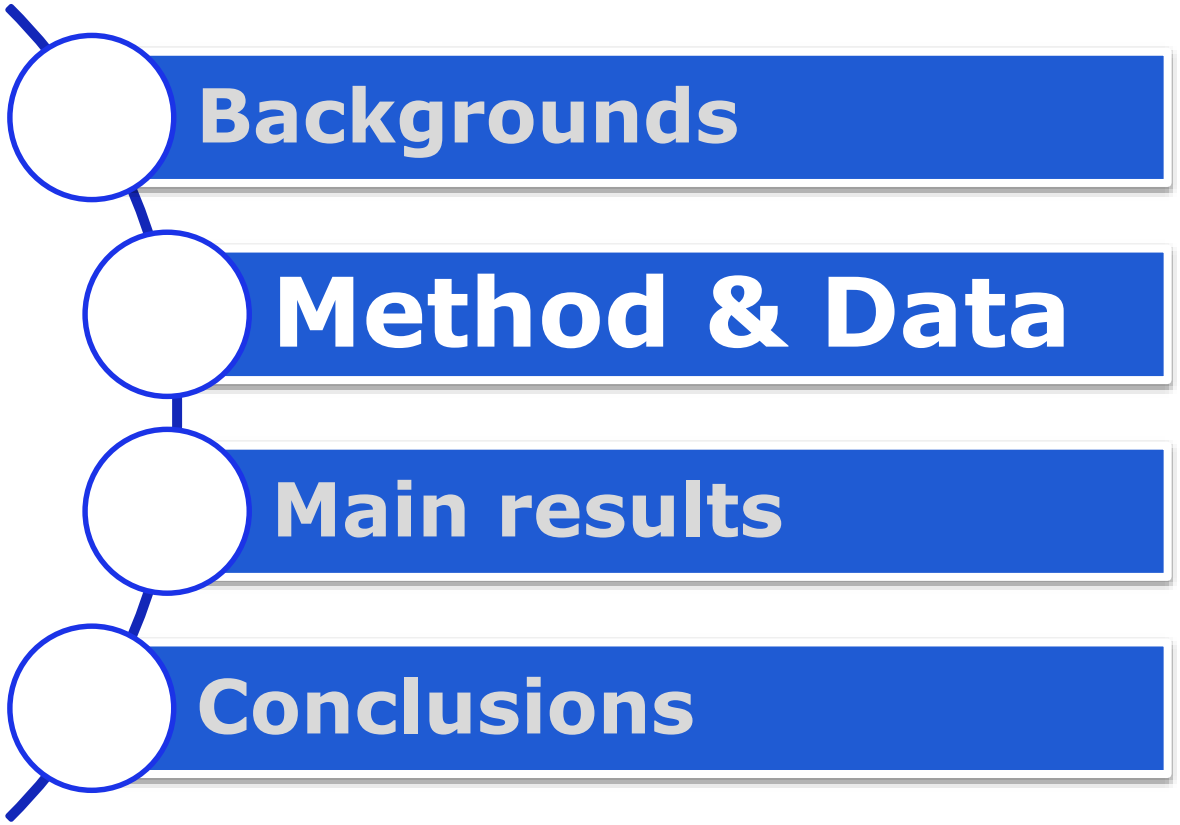
- A promising way
  - To build reputation systems based on users' grouping behavior

Muchnik L., Aral S. and Taylor S.-J., Science, 341 (2013) 647.

Wang T., Wang D. and Wang F., ACM SIGKDD. 2014. 1087-1096.

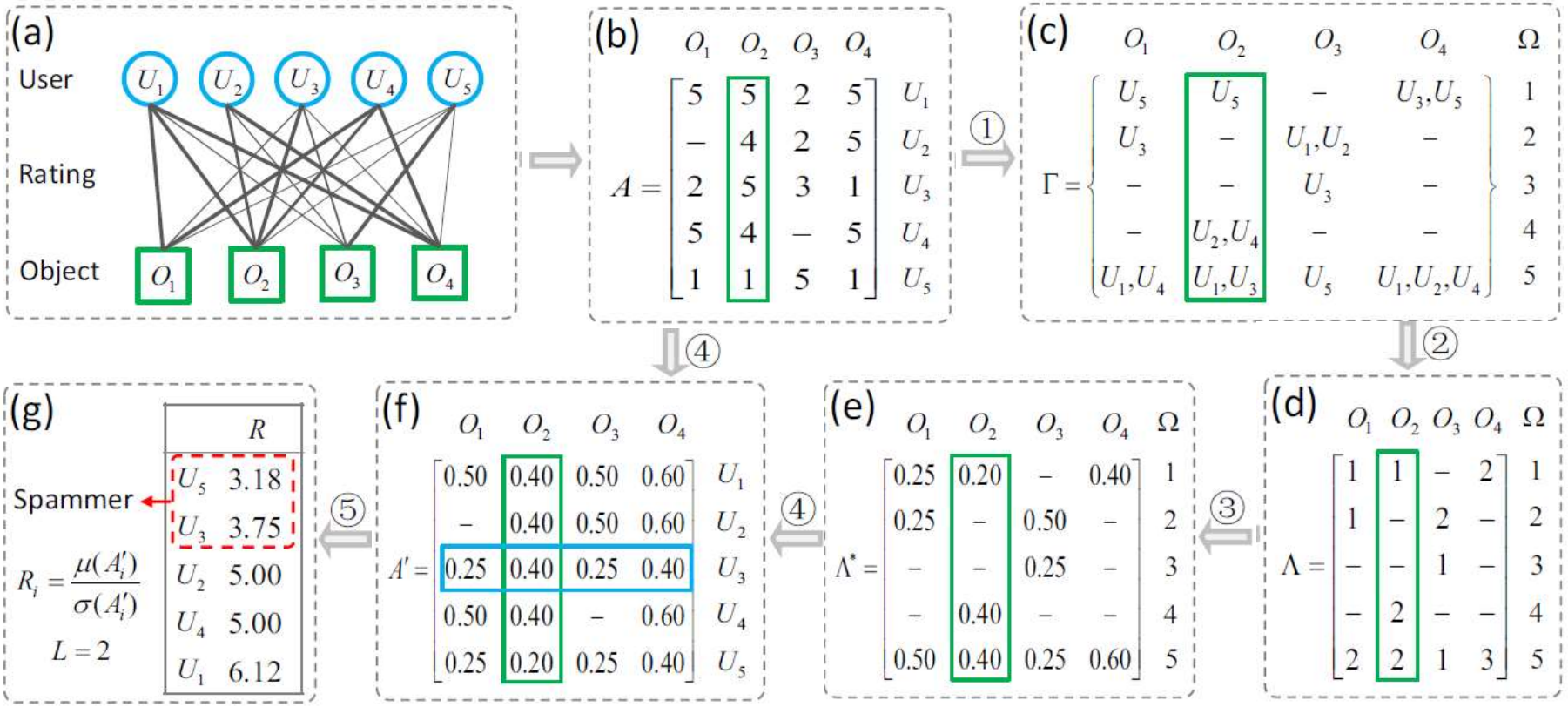
Wang T. and Wang D., Big Data. 2 (2014). 196-204

Outline



Method & Data

- Group-based Ranking (GR) method



Method & Data

- Group-based Ranking (GR) method

$$G = \{U, O, E\}$$

$$U = \{U_1, U_2, ..., U_m\}$$

$$O = \{O_1, O_2, ..., O_n\}$$

$$E = \{E_1, E_2, ..., E_l\}$$

$$a_{i\alpha} \in \Omega = \{\omega_1, \omega_2, ..., \omega_z\}$$

$$\Gamma_{s\alpha} = \{U_i \mid a_{i\alpha} = \omega_s, i = 1, 2, ..., m\}. \quad (1)$$

$$\Lambda_{s\alpha} = |\Gamma_{s\alpha}|$$

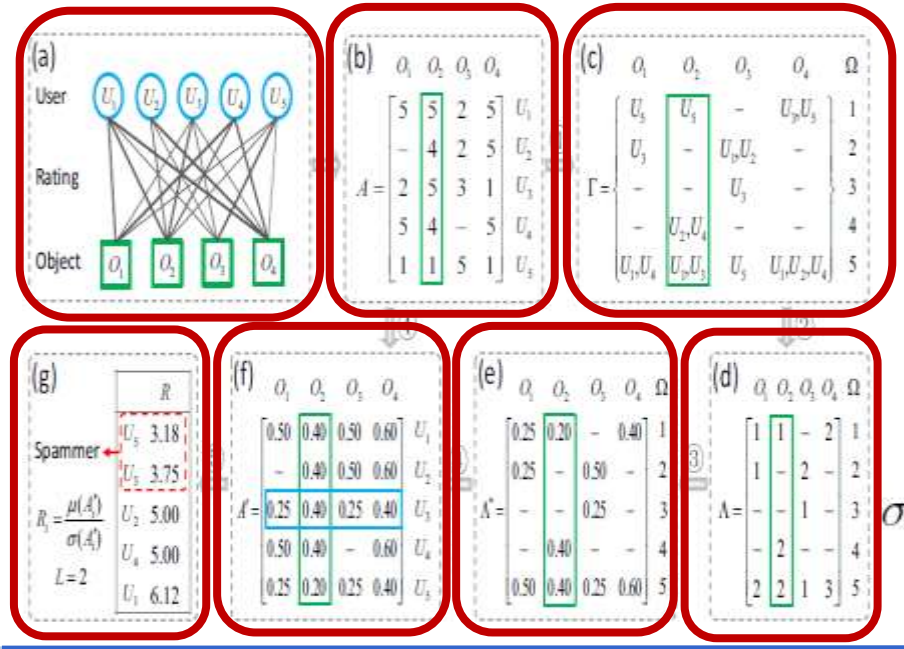
$$\Lambda^*_{s\alpha} = \frac{\Lambda_{s\alpha}}{k_\alpha}. \quad (2)$$

$$A'_{i\alpha} = \Lambda^*_{s\alpha}, \text{ where } a_{i\alpha} = \omega_s$$

$$R_i = \frac{\mu(A'_i)}{\sigma(A'_i)}, \quad (3)$$

$$\mu(A'_i) = \sum_\alpha \frac{A'_{i\alpha}}{k_i}, \quad (4)$$

$$\sigma(A'_i) = \sqrt{\frac{\sum_\alpha (A'_{i\alpha} - \mu(A'_i))^2}{k_i}}. \quad (5)$$



Method & Data

- Data sets

Data set	m	n	$\langle k_U \rangle$	$\langle k_O \rangle$	S
MovieLens	943	1682	106	60	0.063
Netflix	1038	1215	47	40	0.039
Amazon	662	1500	36	15	0.023

- Metrics
  - Recall:

$$R_c(L) = \frac{d'(L)}{d}, \tag{6}$$

  - AUC:

$$AUC = \frac{N' + 0.5N''}{N}. \tag{7}$$

Method & Data

- Spamming behaviors
  - Malicious spammers: minimum (maximum) allowable ratings to push down (up)
  - Random spammers: random ratings, which is meaningless
- Generating artificial spammers
  - Malicious spammers: integer 1 or 5 with the same probability (i.e., 0.5)
  - Random spammers: random integers in {1, 2, 3, 4, 5}
- Ratio and activity is $q = d/m$ and $p = k/n$, respectively

Outline

- Backgrounds
- Method & Data
- **Main results**
- Conclusions

Main results

- Rating error of users

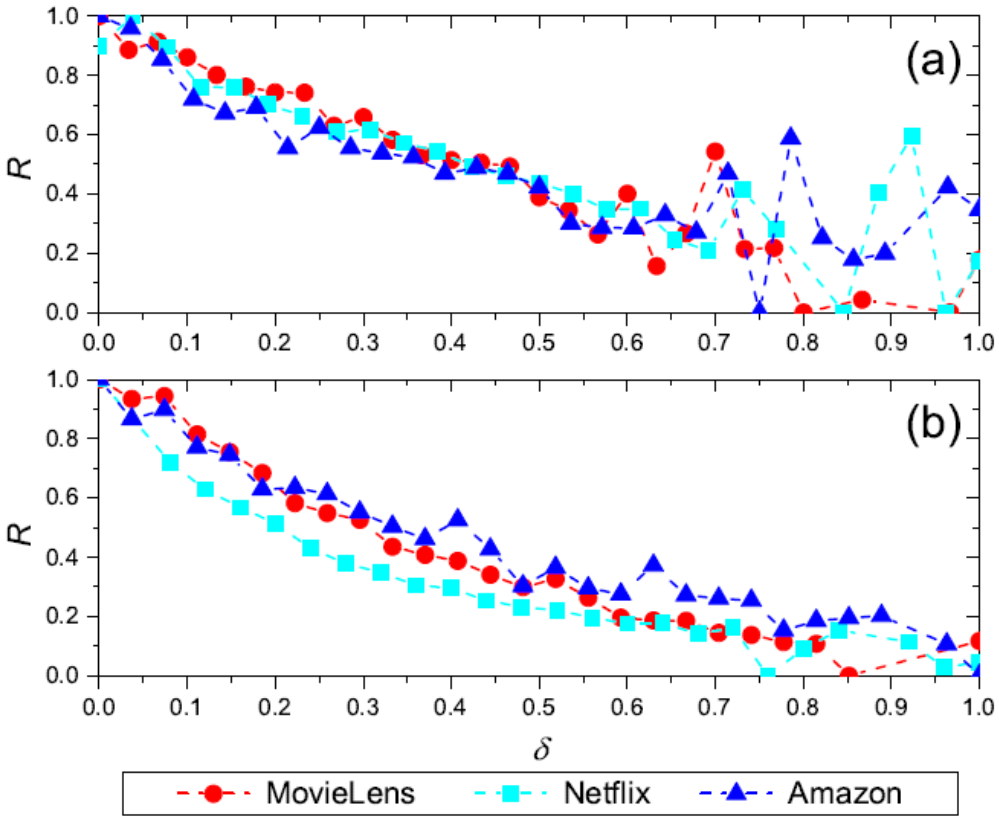
$$\delta_i = \frac{\sum_\alpha |a_{i\alpha} - \hat{Q}_\alpha|}{k_i}$$

$$\hat{Q}_\alpha = \sum_{j \in \Gamma_\alpha} a_{j\alpha}/k_\alpha$$

- ρ = −0.956 (−0.949), −0.906 (−0.872) and −0.966 (−0.816) for GR (CR)

- GR method is better than CR on evaluating users' reputations

CR method



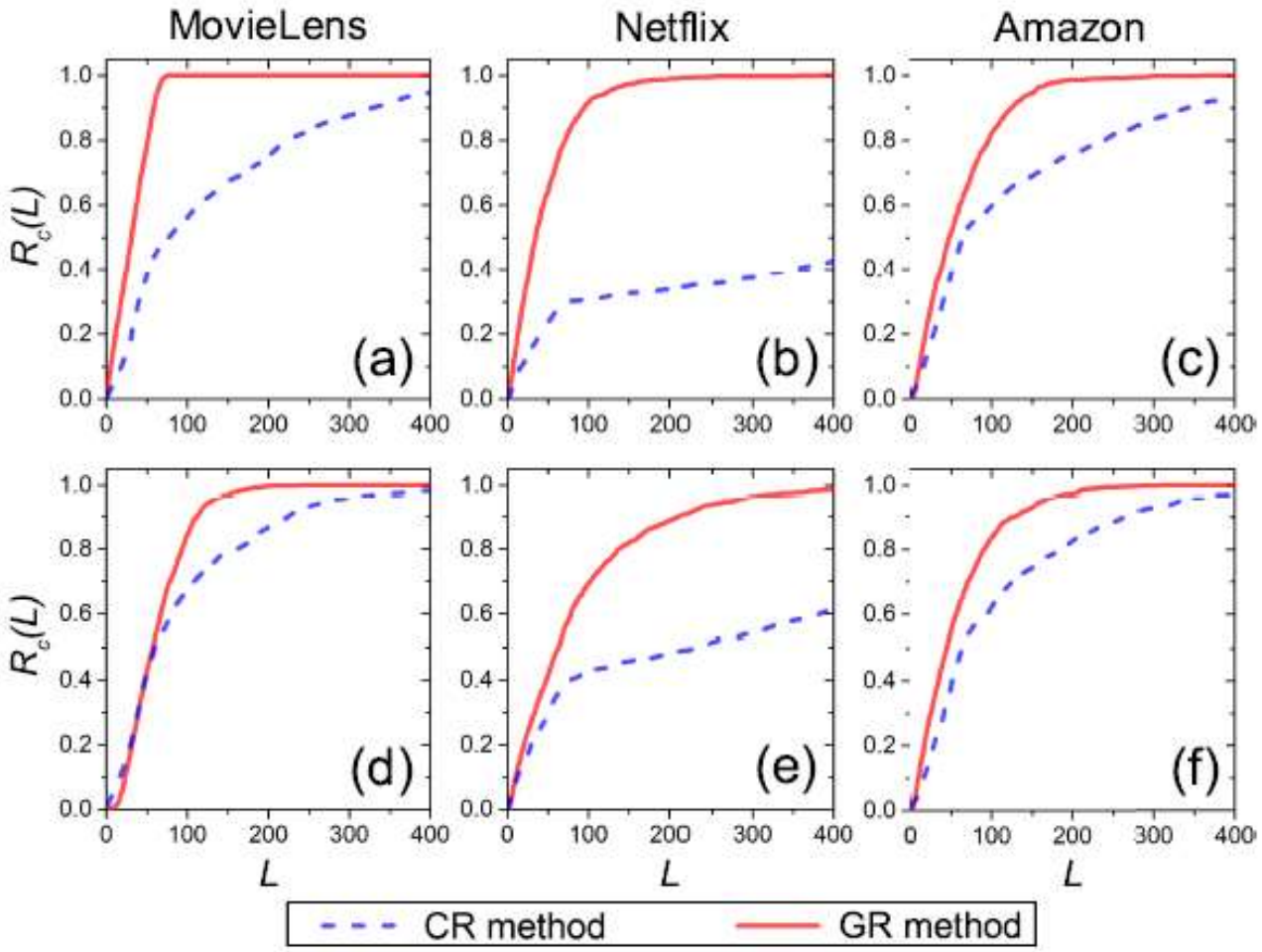
GR method

Main results

- Effectiveness and efficiency

Malicious

Random



Main results

- Robustness against spammers

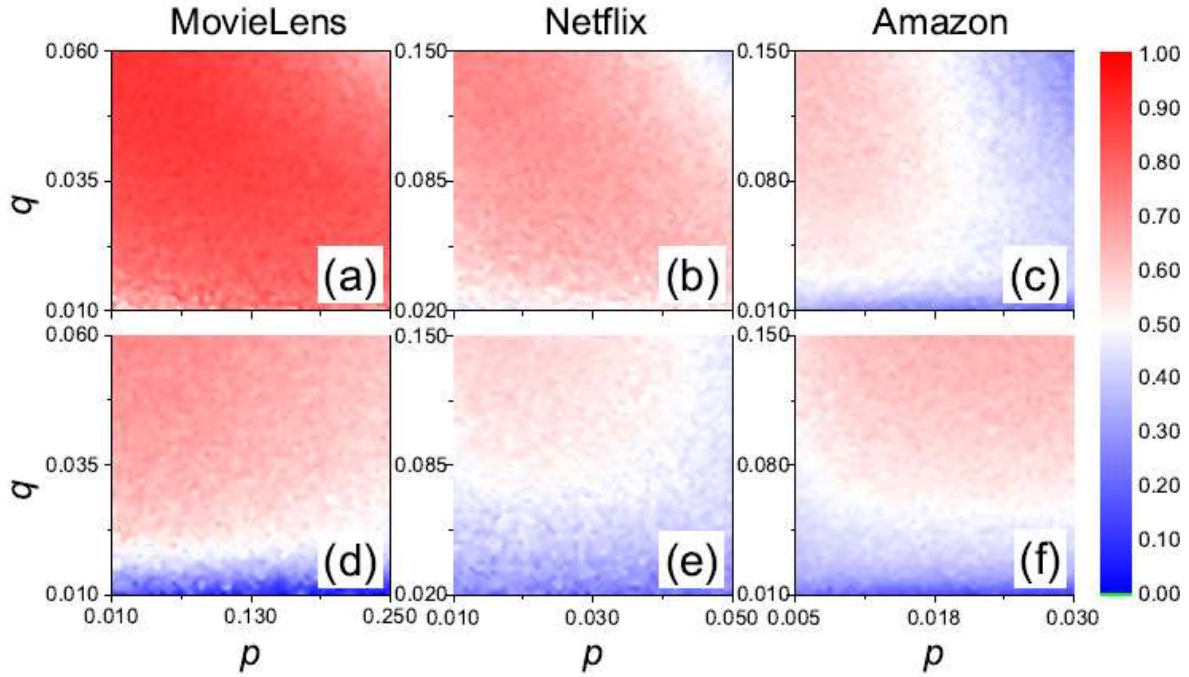


Fig. 4: (Color online) The effectiveness of GR method. The color marks recall R_c. q and p are ratio of spammers and ratio of objects rated by spammers, respectively. (a), (b) and (c) are for malicious spammers. (d), (e) and (f) are for random spammers. The parameter is set as $L = d$. The results are averaged over 100 independent realizations.

Main results

- Robustness against spammers

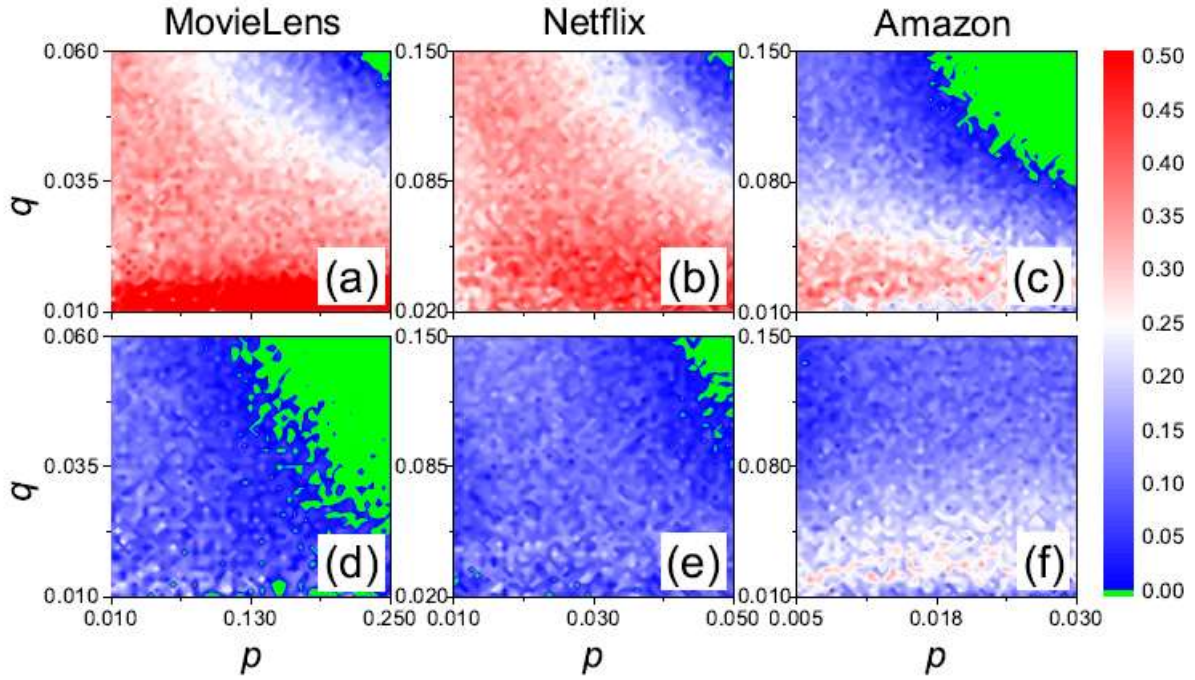


Fig. 5: (Color online) The comparison of GR and CR methods. The color marks ΔR_c if $\Delta R_c > 0$, otherwise the color is green, meaning that the CR method is better. (a), (b) and (c) are for malicious spammers. (d), (e) and (f) are for random spammers. The parameter is set as $L = d$. The results are averaged over 100 independent realizations.

Outline

Conclusions

- We proposed a group-based ranking (GR) method that evaluates users' reputations based on their grouping behaviors for ORS with spamming attacks.
- In testing with three real data sets, GR method gives higher performance in both accuracy and robustness than the CR method.
- The GR method is very efficient, as its time complexity is $O(m^2)$, which is significantly lower than most of previously proposed iterative methods.

Thanks for your attention!

Question or Comment?

JIAN GAO

gaojian08@hotmail.com

CompleX Lab, Web Sciences Center, UESTC

March 28, 2015