

COMMUNITY INFLUENCE ANALYSIS IN SOCIAL NETWORK

Ximing Li

AGENDA

- ▶ Introduction
- ▶ Motivation
- ▶ Problem Definition
- ▶ Algorithm
- ▶ Implementation
- ▶ Conclusion&Future work

INTRODUCTION

- ▶ Background:
 - ▶ Social Network
 - ▶ Social influence

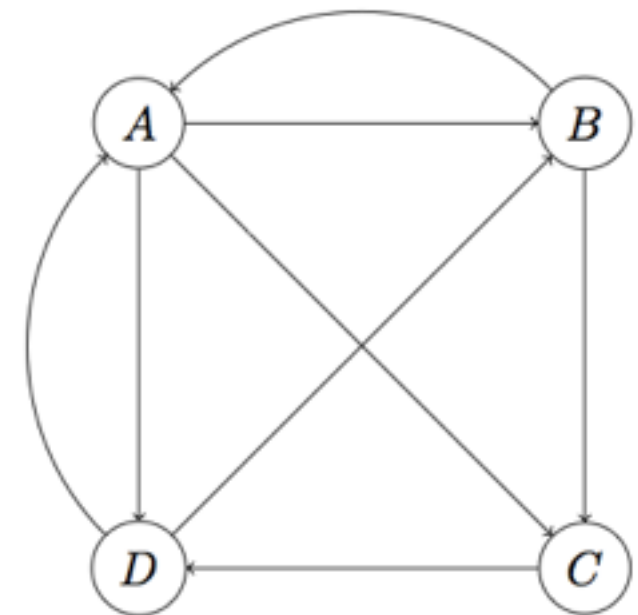


MOTIVATION

- ▶ Community-level influence
- ▶ Directed graph with weightage
- ▶ Internal and external influence
- ▶ Centrality information

PROBLEM & DEFINITION

- ▶ Community-Level influence
- ▶ Citation Network: $G\{V,E\}$
- ▶ Joint weight based directed graph
 - ▶ V : node(Venue)
 - ▶ E : edge associate with weightage
- ▶ Identifying the most influential community



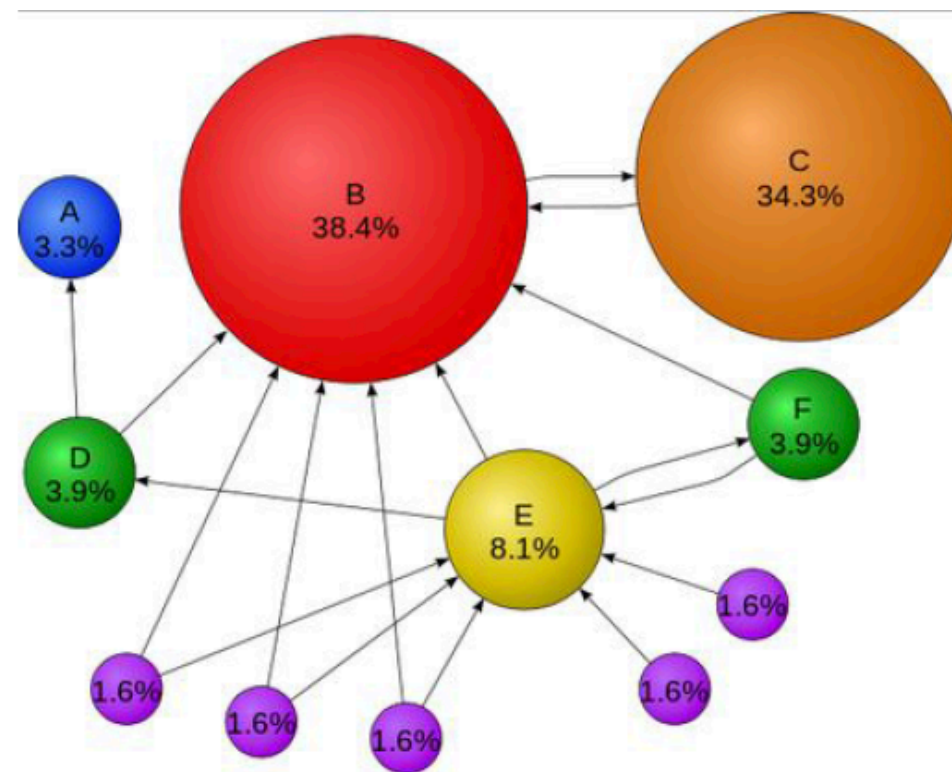
ALGORITHM-PAGERANK

- ▶ Page Rank is "vote" by all the other nodes.
- ▶ $PR(A) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$

Algorithm outline:

- Randomly assign vector X with positive numbers, and the length of X is same as the number of nodes
- Repeat
 - $X = R$
 - $R = S * X$
- Until $|X-R| < \theta$ // θ is threshold value

Where S is the source matrix of graph



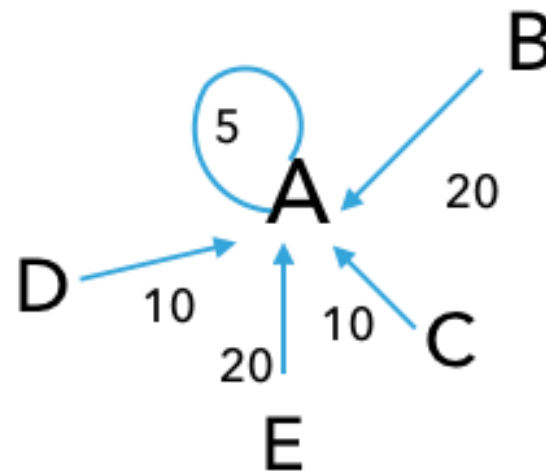
ALGORITHM MODIFY

► Modify

► Ignore internal influence

► Ignore weight

► Wight formula: $w_i = \sum_{i=1}^n \sum_{j=1}^n \left(\frac{c_{ij}}{\left(\frac{T_i + T_j}{2} \right)} * C_{id} \right)$



ALGORITHM DESIGN

► ComRank:

Algorithm: ComRank

Input: weight matrix w and $G(V,E)$

Output: Rank vector R

1. set *threshold value*
 2. set damping factor d ; //between 0 to 1
 3. $n = |G|$
 4. uniform matrix e^T
 5. for $i = 0$ to $n - 1$
 6. initial value $\rightarrow CR[i]$
 7. transition matrix w'
 8. $R[0] = w' * CR[0]$
 9. for $i = 1$ to n
 10. while $|R[i] - R[i-1]| > \text{threshold value}$ do
 11. $R[i] = R[i-1] * CR[i]$
 12. continue;
 13. if $|R[i] - R[i-1]| < \text{threshold value}$ then
 14. break;
 15. collections.sort vector R .
-

IMPLEMENTATION

- ▶ Dataset: DBLP V7
 - ▶ Paper: 2,244,021
 - ▶ Citation Relationship: 4,354,534
 - ▶ Ranging from 1995 to 2016
- ▶ Each Venue can be seen as a community

```
#*A three-stage approach for the resource-constrained shortest path as a sub-problem  
in column generation.  
#@Xiaoyan Zhu,Wilbert E. Wilhelm  
#t2012  
#cComputers & OR  
#index3063614  
#%186398  
#%811958  
#%885474  
#%924122  
#!
```

IMPLEMENTATION-CITATION RELATIONSHIP

► Citation relationship

```
#cAAAI\\ &&& #cAAAI\\
#cAAAI\\ &&& #cAAMAS (1)\\
#cAAAI\\ &&& #cAAMAS (2)\\
#cAAAI\\ &&& #cAAMAS (3)\\
#cAAAI\\ &&& #cAAMAS (Industry Track)\\
#cAAAI\\ &&& #cAAMAS\\
#cAAAI\\ &&& #cACL/AFNLP (Short Papers)\\
#cAAAI\\ &&& #cACL/AFNLP\\
#cAAAI\\ &&& #cAGI\\
#cAAAI\\ &&& #cARES\\
#cAAAI\\ &&& #cAnn. Math. \\ul Artif\\ulnone . \\ul Intell\\ulnone .\\
#cAAAI\\ &&& #cArtif. \\ul Intell\\ulnone .\\
#cAAAI\\ &&& #cArtificial Life\\
#cAAAI\\ &&& #cAuton. Robots\\
#cAAAI\\ &&& #cAutonomous Agents and \\ul Multi\\ulnone -Agent Systems\\
#cAAAI\\ &&& #cCATS\\
#cAAAI\\ &&& #cCHI Extended Abstracts\\
#cAAAI\\ &&& #cCHI\\
#cAAAI\\ &&& #cCoRR\\
#cAAAI\\ &&& #cCommun. ACM\\
#cAAAI\\ &&& #cComputational Linguistics\\
#cAAAI\\ &&& #cConstraints\\
#cAAAI\\ &&& #cECIR\\
#cAAAI\\ &&& #cECML/PKDD (2)\\
#cAAAI\\ &&& #cHICSS\\
#cAAAI\\ &&& #cHRI\\
```

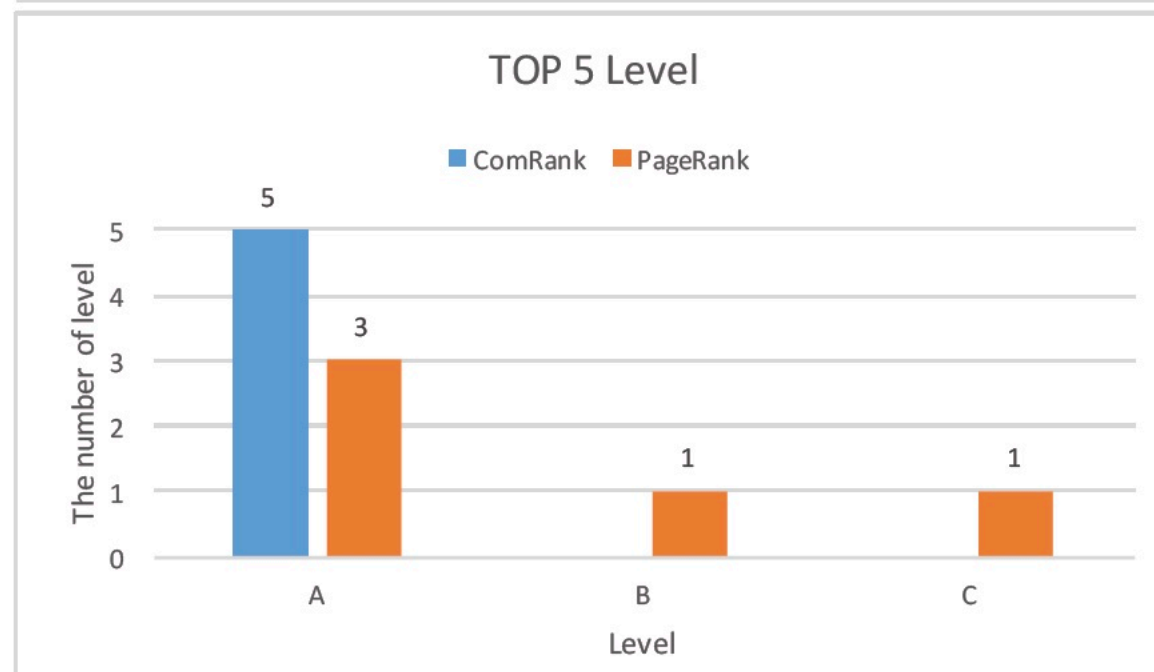
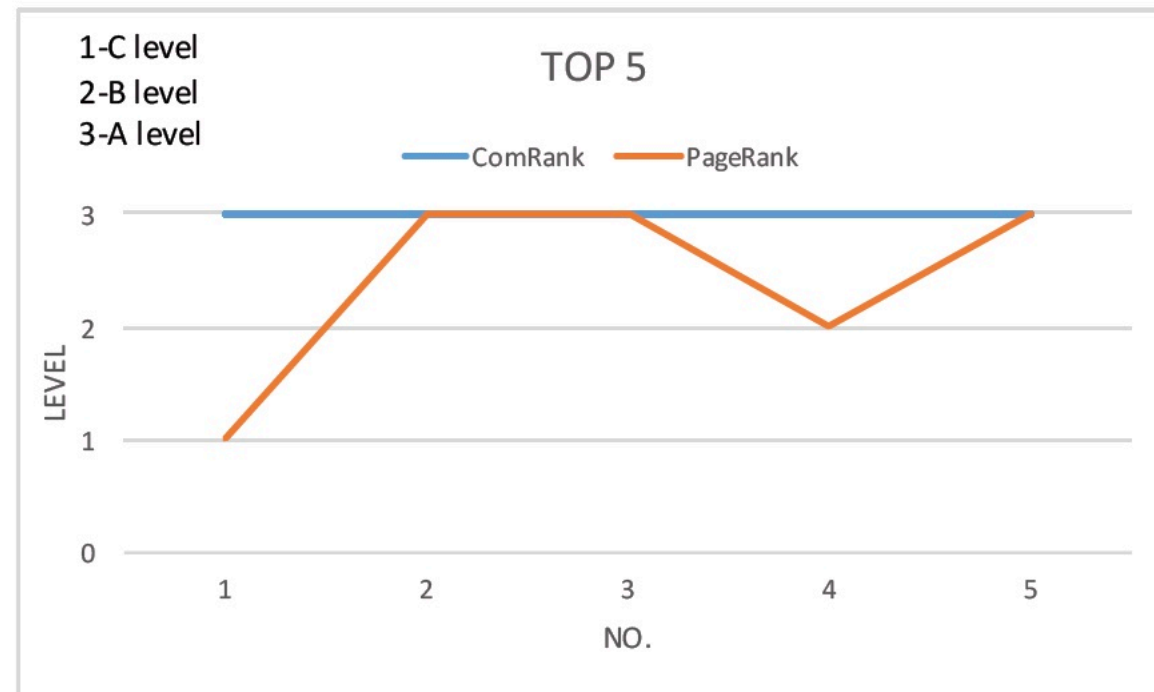
Citation Weight

```
#cAAAI\\ &&& #cAAAI\\ : 0.051960784313725486
#cAAAI\\ &&& #cAAMAS (1)\\ : 0.051960784313725486
#cAAAI\\ &&& #cAAMAS (2)\\ : 0.4676470588235294
#cAAAI\\ &&& #cAAMAS (3)\\ : 0.15588235294117647
#cAAAI\\ &&& #cAAMAS (Industry Track)\\ : 0.051960784313725486
#cAAAI\\ &&& #cAAMAS\\ : 0.051960784313725486
#cAAAI\\ &&& #cACL/AFNLP (Short Papers)\\ : 0.051960784313725486
#cAAAI\\ &&& #cACL/AFNLP\\ : 0.051960784313725486
#cAAAI\\ &&& #cAGI\\ : 0.051960784313725486
#cAAAI\\ &&& #cARES\\ : 0.051960784313725486
#cAAAI\\ &&& #cAnn. Math. \\ul Artif\\ulnone . \\ul Intell\\ulnone .\\ : 0.051960784313725
#cAAAI\\ &&& #cArtif. \\ul Intell\\ulnone .\\ : 0.20784313725490194
#cAAAI\\ &&& #cArtificial Life\\ : 0.051960784313725486
#cAAAI\\ &&& #cAuton. Robots\\ : 0.051960784313725486
#cAAAI\\ &&& #cAutonomous Agents and \\ul Multi\\ulnone -Agent Systems\\ : 0.05196078431
#cAAAI\\ &&& #cCATS\\ : 0.051960784313725486
#cAAAI\\ &&& #cCHI Extended Abstracts\\ : 0.051960784313725486
#cAAAI\\ &&& #cCHI\\ : 0.20784313725490194
#cAAAI\\ &&& #cCoRR\\ : 0.051960784313725486
#cAAAI\\ &&& #cCommun. ACM\\ : 0.051960784313725486
#cAAAI\\ &&& #cComputational Linguistics\\ : 0.051960784313725486
#cAAAI\\ &&& #cConstraints\\ : 0.051960784313725486
#cAAAI\\ &&& #cECIR\\ : 0.051960784313725486
#cAAAI\\ &&& #cECML/PKDD (2)\\ : 0.051960784313725486
#cAAAI\\ &&& #cHICSS\\ : 0.051960784313725486
#cAAAI\\ &&& #cHRI\\ : 0.10392156862745097
#cAAAI\\ &&& #cI. J. Robotic Res.\\ : 0.15588235294117647
#cAAAI\\ &&& #cIAT\\ : 0.051960784313725486
#cAAAI\\ &&& #cICALT\\ : 0.051960784313725486
```

IMPLEMENTATION-RESULT TOP5

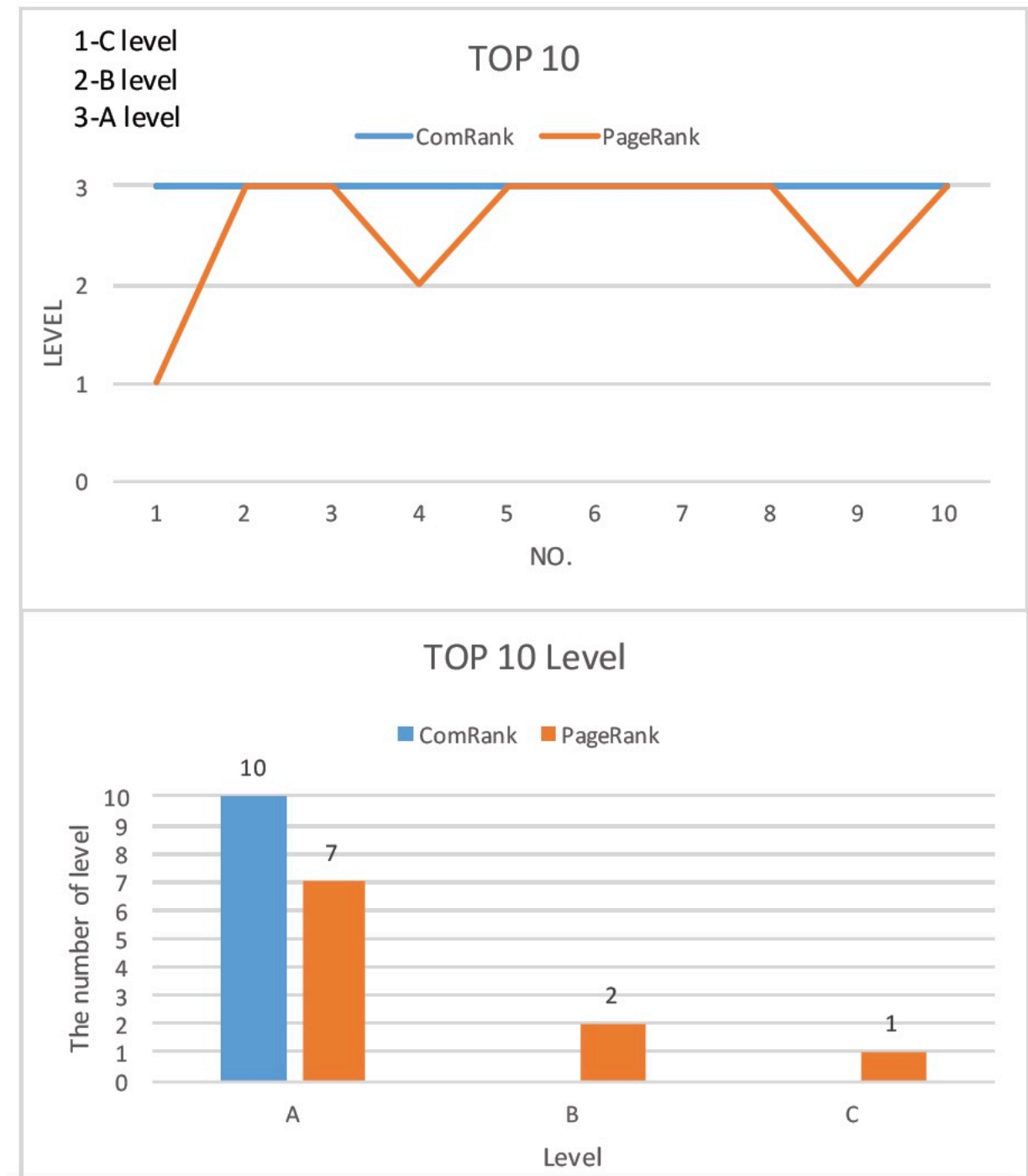
► Community Ranking:

Rank	ComRank		PageRank	
NO.	Community	Level	Community	Level
1	AAAI	A	SAC	C
2	ICDE	A	ICML	A
3	SIGCOMM	A	IJCAI	A
4	CRYPTO	A	GECCO	B
5	KDD	A	AAAI	A



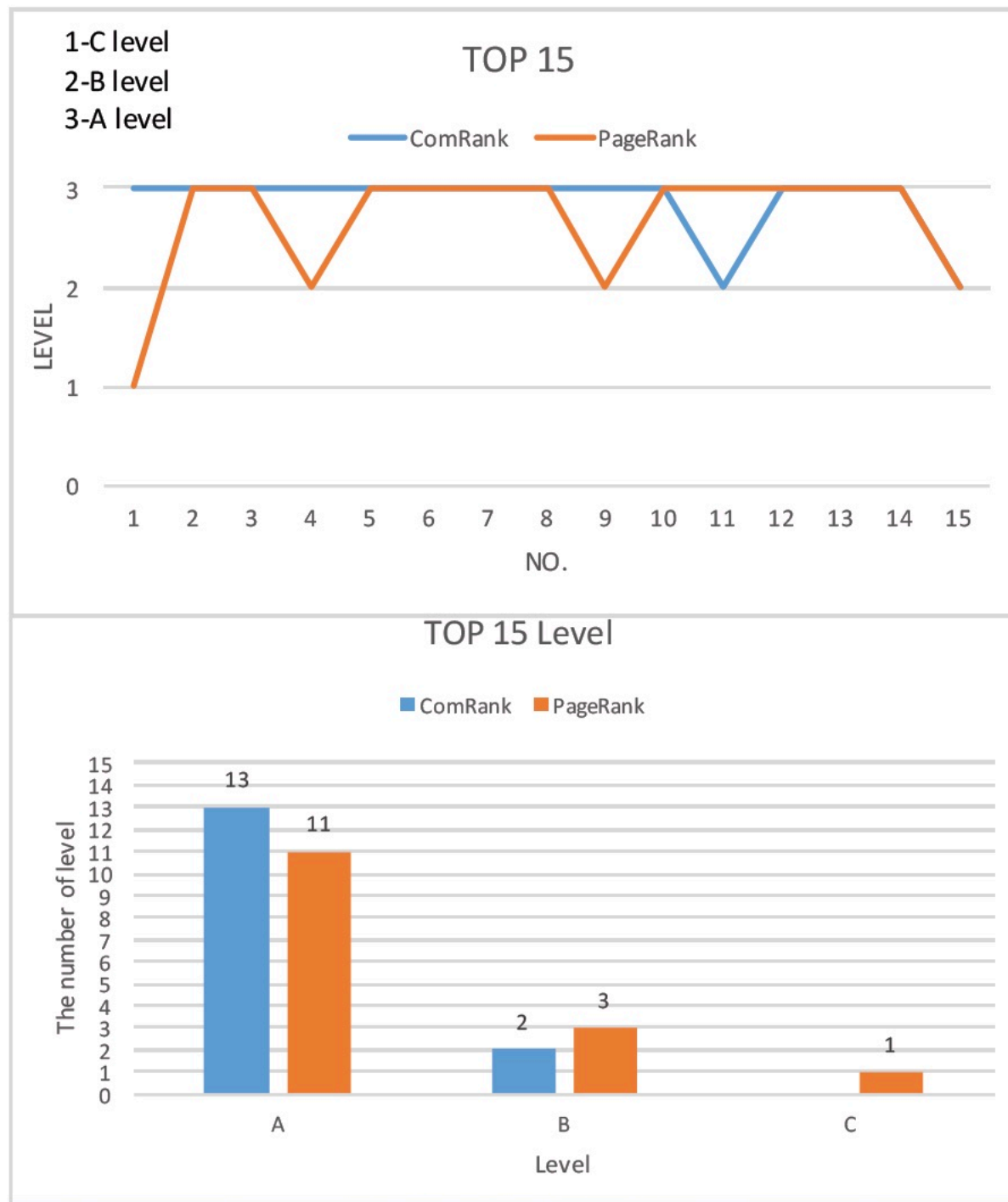
IMPLEMENTATION-RESULT TOP10

Rank	ComRank		PageRank	
NO.	Community	Level	Community	Level
1	AAAI	A	SAC	C
2	ICDE	A	ICML	A
3	SIGCOMM	A	IJCAI	A
4	CRYPTO	A	GECCO	B
5	KDD	A	AAAI	A
6	SODA	A	SODA	A
7	SIGIR	A	WWW	A
8	CIKM	A	SIGCOMM	A
9	ICCV	A	EMNLP	B
10	ICCAD	A	STOC	A



IMPLEMENTATION-RESULT TOP15

Rank	ComRank		PageRank	
NO.	Community	Level	Community	Level
1	AAAI	A	SAC	C
2	ICDE	A	ICML	A
3	SIGCOMM	A	IJCAI	A
4	CRYPTO	A	GECCO	B
5	KDD	A	AAAI	A
6	SODA	A	SODA	A
7	SIGIR	A	WWW	A
8	CIKM	A	SIGCOMM	A
9	ICCV	A	EMNLP	B
10	ICCAD	A	STOC	A
11	TACAS	B	CRYPTO	A
12	IJCAI	A	KDD	A
13	ICML	A	SIGIR	A
14	FOCS	A	ICDE	A
15	WSDM	B	WSDM	B



CONCLUSION & FUTURE WORK

- ▶ Conclusion
 - ▶ Community-level influence analysis
 - ▶ Assign weight based on joint method
 - ▶ Internal influence of community
 - ▶ Modify PageRank algorithm to incorporate weight
- ▶ Future Work
 - ▶ Apply cluster to generate community
 - ▶ More factor, more data
 - ▶ Citation overlapping

REFERENCE

- [1]. *Erjia Yan, Ying Ding, Discovering author impact: A PageRank perspective, in Information Processing and Management 47 125-134*
- [2]. *Citation Network Dataset, available form: <https://aminer.org/billboard/citation>*
- [3]. *Conference level, available form: <http://www.ntu.edu.sg/home/assourav/crank.htm>.*
- [4]. *Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. ArnetMiner: Extraction and Mining of Academic Social Networks. In Proceedings of the Fourteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD). pp.990-998.*

THANK YOU!

ALGORITHM MODIFY

- ▶ Different edge has different weight.
- ▶ Weight formula 1: $\frac{A \text{ got citation form } B}{A \text{ total citaton}}$
- ▶ Weight formula 2: $\frac{A \text{ got citation form } B}{B \text{ total citaton}}$