

ָפָ ר a פָ , ר a ... פָ a פָ , `` , h... , Member, IEEEa , a פָ , Senior Member, IEEE

Abstract-ΪĜ а ŧ/ ; , / <sub>.1</sub> H រេឡូជូ "a ∣ а a t h t<sub>∥</sub>at а ΪĜ <sub>I .I</sub>∉ a ١I ъđ 11 1 h Qh .fı'i h 24 ta ħ аa t f d. ١. L J Т 1 Hata jath ĝh ″" ""â a tt 1 ht tta а а а ١, d. AL. а<sub>I</sub> h ĝh a at Ht **⊈** đi <sub>I</sub> a а ų I.I. L d ١. J. ŀ "fi "j**á** t, а ,₁a t<sub>i</sub> а , ,at, а a t<sub>i</sub> а а Ja d. a Ha t  $\boldsymbol{t}_{d}$ at¦₿ ∣åt ļ Ηt. , a t tt а t, Jfia ţ ί. Ы а ١ t<sub>i</sub>a \ \_ ta i H a ti <sub>l</sub>a а a a 1 ′at<sub>l</sub> **h** ł t h ī a ١I Цų

Index Terms— 1 111, ta fat, a t

#### **1** INTRODUCTION



- J. Zhang and Z. Fang are with the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China. E-mail: {zhangjing12, fzp13}@mails.tsinghua.edu.cn.
- W. Chen is with Theory Group, Microsoft Research, Beijing 100080, China. E-mail: weic@microsoft.com.
- J. Tang is with the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China, and Tsinghua National Laboratory for Information Science and Technology (TNList). E-mail: jietang@tsinghua.edu.cn.

Manuscript received 11 Jan. 2014; revised 5 Jan. 2015; accepted 9 Feb. 2015. Date of publication 25 Feb. 2015; date of current version 2 July 2015. Recommended for acceptance by G. Das.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TKDE.2015.2407351

0,

).

(...,

).

A

",

)

(+ ) (+ ) "following" link cascade model.

.)

Organization.

· , . . . .

# 2 "FOLLOWING" LINK CASCADE MODEL

 $t(e_{uv}) = n \in e_{uv} \qquad n,$   $t(e_{uv}) = \bot$   $e_{uv} \qquad ,$   $t(e_{uv}) = \bot$   $C, e \qquad B \quad C.$   $A \qquad 1. Diffusion effect between links decays over time.$ 

discovery probability  $g_{e'e}$ BACdiffusion probability  $h_{e'e}$ CA В C. t', t'e  $t' + \delta$ ,  $g_{e'e'}_{e'}$ e'















sion.	A			C
				B
$\leq t \leq t' + \delta$ ,		B	A'	
' - 1,				
$e_{AC}$				$e_{BC}$ .
			C	
	sion. $t \le t \le t' + \delta,$ t' - 1, $e_{AC}$	sion. A $t \leq t \leq t' + \delta,$ t' - 1, $e_{AC}$	sion. A $t \leq t \leq t' + \delta, \qquad B$ $t'-1, \qquad e_{AC}$	sion. A $t \leq t \leq t' + \delta$ , B A' t' - 1, $e_{AC}$ C

follower diffusion.



, . ., е<sub>АС</sub> е<sub>СВ</sub>.

#### **3** DATA AND OBSERVATIONS

C

'

# 3.1 Data Collection

, , . ., , 0,000 " ' / / 0 0. 0/ / 0 0 , 11 ,0 ., 0 0/ / 0 0 / / 0 0. ( )

# 3.2 Observations

()

1 12 . 13 24 ( )

( ) t'ť, t

 $0 \leq t - t' \leq \delta$  ( $\delta$ ).  $C_{\bigtriangleup}$ ,  $\begin{array}{c} \overleftarrow{\phantom{a}}, & C^+_{\Delta} \\ [t', t' + \delta]. & |C_{\Delta}| \\ \Delta. \end{array}$ riangle,CB $r_{\triangle}$ CB

$$r_{\triangle} = \frac{|C_{\triangle}^+|}{|C_{\triangle}|}.\tag{1}$$

Pattern significance.

. ,





. , δ





-

# 4 MODEL LEARNING

$$heta = \{h_{e'e}, g_{e'e}\}$$
 -

(e',e).

$$g_{e'e} \qquad (e',e)$$
 (  $e$ 

$$h_{e'e}$$

$$e \in \mathcal{E}$$

$$\log \mathcal{L} = \sum_{e \in \mathcal{E}} \left\{ \log \sum_{\vec{\alpha}_{S_e}} \prod_{e' \in S_e} x_{e'e}^{\alpha_{e'}} y_{e'e}^{1-\alpha_{e'}} + \sum_{e' \in R_e} \log y_{ee'} \right\}.$$

EM algorithm. .

-

$$q(e|\vec{\alpha}_{S_e}) = \frac{p(e|\vec{\alpha}_{S_e})}{p(e|S_e)}$$

\_

.

$$\log \mathcal{L} = \sum_{e \in \mathcal{E}} \left\{ \log \sum_{\vec{\alpha}_{S_e}} \hat{q}(e | \vec{\alpha}_{S_e}) \frac{p(e | \vec{\alpha}_{S_e})}{\hat{q}(e | \vec{\alpha}_{S_e})} + \sum_{e' \in R_e} \log y_{ee'} \right\}$$
$$\geq \sum_{e \in \mathcal{E}} \left\{ \sum_{\vec{\alpha}_{S_e}} \hat{q}(e | \vec{\alpha}_{S_e}) \log \frac{p(e | \vec{\alpha}_{S_e})}{\hat{q}(e | \vec{\alpha}_{S_e})} + \sum_{e' \in R_e} \log y_{ee'} \right\},$$
$$\hat{q}(e | \vec{\alpha}_{S_e}) \log \hat{q}(e | \vec{\alpha}_{S_e})$$

,

 $Q(\theta, \hat{\theta})$ 

$$Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \sum_{\boldsymbol{e} \in \mathcal{E}} \left\{ \sum_{\vec{\alpha}_S} \right.$$

2100

. I. I. I. I. .

v

$$h_{\Delta} = \frac{\sum_{(e',e) \in C_{\Delta}^{+}} \hat{D}_{e'e} + \sum_{(e',e) \in C_{\Delta}^{-}} \hat{B}_{e'e}}{|C_{\Delta}|},$$
(12)

1

$$g_{\triangle} = \frac{\sum_{(e',e)\in C_{\triangle}^{+}} A_{e'e}}{\sum_{(e',e)\in C_{\triangle}^{-}} \hat{B}_{ee'}(\delta+1) + \sum_{(e',e)\in C_{\triangle}^{+}} \hat{D}_{e'e}(t_e - t_{e'} + 1)}.$$
 (13)

$$() (), C_{\Delta}^{+} C_{\Delta}^{-} .$$

$$A_{e'e}, B_{e'e} D_{e'e}$$

$$(), D_{e'e} B_{e'e} - A_{e'e} B_{e'e}$$

$$D_{e'e} = B_{e'e} + A_{e'e} - A_{e'e} B_{e'e}.$$

$$(14)$$

v .0. -

#### Followee maximization.

A - 2.	/			
:	G = (V, E),	ν,	k	
:	/	S		
$S = \emptyset$	<i>R</i> - 0,000			
$i = 1 \ to \ k\mathbf{q}$				
$\neg u \in V$	\S <b>d</b>			
$s_u = 0$				
r = 1 to	R			
$s_u + =  FC $	$CM(S \cup \{u\}) $			
$s_u = s_u/R$				
$S = S \cup \{argn$	$nax_{u \in V \setminus S} s_u$			

### **6 EXPERIMENTS**

,

## **5 APPLICATIONS**

, Follower maximization. ν, kSvi $u \notin S$ , v.  $(S \cup \{u\})$  ( ), (S)S.( . ), Sve'. ( ).

е *S*,

## 6.1 Experimental Setup



· / /

. Evaluation metrics.

# 

 $p(e|S_e) ,$   $p(e|S_e) ,$  e ,  $p(e|S_e) > \tau. ,$   $p(e|S_e) > \tau. ,$ 

 $p(e|S_e)$  -

#### Comparison methods.

Basic.

.

LRC.

. u w, u

2101

$$CF\_score(u,v) = \sum_{w} I(w,v) sim(w,u),$$

и,

 $\{g_{\triangle}\}$ 

,

sim(w, u)wи, I(w, v)0 wv\_  $\{v\}$ u $CF\_score(u, v).$ SimRank. uv0.  $\{v\}$ uKatz. uvuv.

#### Random-random model (RR). - . , u v, u v, u w v, u w v, u v e\_uv - .

(u,w,v) . , u v

$$RR\_score(u,v) = \frac{1}{|F(u)|} \sum_{w} I(u,w)I(w,v)\frac{1}{|F(w)|},$$

Preferential attachment with communities (PAC).

и,	0.V				
и,					
<i>u</i> .	α	β	0	0	-
		1			-



#### М alysis. lel parame ei а

.1, 7

kv. vv.( -

). 0.0 . k/ , . 0 / ).

(+ -.0

/ . . .

-

\_

, –

# 7 RELATED WORK

Diffusion model and influence maximization.

•

,

- Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00,
- Proc. 4th ACM Int. Conf. Web Search Data Mining, '0,
- , Science, . , . 0 . , 111.
- ing, 0 , . **?**0. , Proc. IEEE 12th Int. Conf. Data Min-

- Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 0 0, . 0 0 .
- Proc. 15thor CM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00, 100
- . , . , . , . , . , . , . , . , . , . Proc. 16th Int. Conf. World Wide Web, 00 , . 0.
- , ", Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00 ,
- , Networks, Crowds, and Markets: Reasoning about a Highly Connected World.
- , Tsinghua Sci. Technol.,
- Search Data Mining, 0 0, . Proc. 3rd ACM Int. Conf. Web
- ", Proc. 13th Int. Conf. World Wide Web, 00, . . 0.
- ing Workshop, 00 .
- 0 . , ", Proc. 8th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00 , ...
- Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00,
- , Proc. Roy. Soc. A, . . , . 00
- Knowl. Manage., 00, . 0 . Proc. ACM Conf. Inf.

- · , · , · , · , ",
- , Intell. Data Anal., . . , . . , 0 .
- , Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 0 ,
- ", Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00, 0.
- Conf. World Wide Web, 0 0,
   0.
   "
   Proc. 21st Int.
- -Proc. 11th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 00, .

- . 0, . **1** , **111**.
- Proc. ACM Int. Conf. Web Search
- Int. Conf. Knowl. Discovery Data Mining, 00,

- Proc. 4th Int. AAAI Conf. Weblogs Social Media, 0 0, .

- ing, 00, . . . Proc. 7th IEEE Int. Conf. Data Min-

**?**...,..,..,...,