Regionally Influential Users in Location-Aware Social Networks

Panagiotis Bouros[1] Dimitris Sacharidis[2] Nikos Bikakis[3]

[1]Humboldt-Universität zu Berlin, Germany
bourospa@informatik.hu-berlin.de
[2]Institute for the Management of Information Systems – "Athena" R.C, Greece
dsachar@imis.athena-innovation.gr
[3]National Technical University of Athens, Greece
bikakis@dblab.ntua.gr

22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems
November 4-7, 2014, Dallas, Texas, USA

1. Motivation

- Motivated by word-of-mouth and viral marketing
- Most influential users within a spatial region
  - Best people to spread the word and raise largest possible attention

2. Problem Definition

Location-Aware Social Network

- Set of users U, set of locations L, set of check-ins C, social graph $G(U,E)$

Propagation model MIAwoT

- Propagation probability
  - p_{xy} for edge (u_x, u_y) of social graph G, degree of influence
  - $p(\pi_{xy})$ for path $\pi_{xy}(u_x, \dots u_y)$ on social graph
- User u_x influences u_y only via maximum influence path (mip) π^*_{xy}

$$p(\pi^*_{xy}) = max_{\forall \pi_{xy} \in G}\{p(\pi_{xy})\}$$

Regional users U_R

- Set of users checked-in at a location inside spatial region R

Locality γ_R

- Probability of checking-in inside region R

$$\gamma_R(u_x) = \frac{|C(u_x) \text{ inside } R|}{|C(u_x)|}$$

Regional influence I_R

- Likelihood of influencing regional users

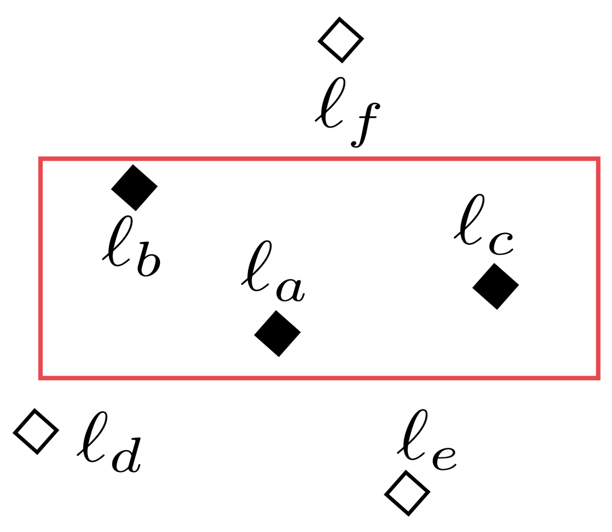
$$I_R(u_x) = \sum_{\forall u_i \in U_R} p(\pi^*_{xi}) \cdot \gamma_R(u_i)$$

Problem k-RIL

Find subset of k regional users $\mathcal{T} \subseteq U_R$: $\forall u_i \in \mathcal{T}$ and $\forall u_j \in U_R \setminus \mathcal{T}$, $I_R(u_i) \geq I_R(u_j)$

Example

Locations L



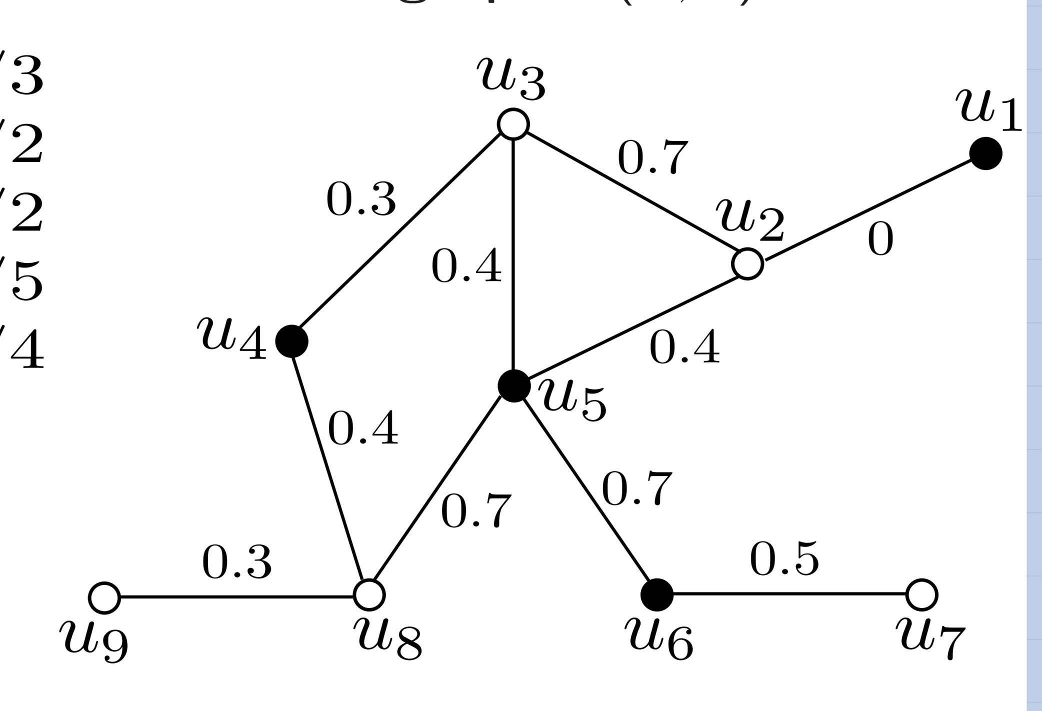
Check-ins C

ℓ_α:	u_1, u_4	R
ℓ_b:	u_1, u_5	
ℓ_c:	u_1, u_5, u_6	
ℓ_d:	u_3, u_9	
ℓ_e:	u_2, u_7, u_8	
ℓ_f:	u_3	

Propagation probabilities

$p_{12} = 1$	$p_{48} = 2/3$
$p_{23} = 1/2$	$p_{56} = 1/2$
$p_{25} = 2/3$	$p_{58} = 1/2$
$p_{34} = 3/4$	$p_{67} = 3/5$
$p_{35} = 2/3$	$p_{89} = 3/4$

Social graph $G(U,E)$



- Regional users, $U_R = \{u_1, u_4, u_5, u_6\}$
- Propagation model
- mips for u_1

$$\pi^*_{11}(u_1),\ \pi^*_{14}(u_1,u_2,u_3,u_4),\ \pi^*_{15}(u_1,u_2,u_5),\ \pi^*_{16}(u_1,u_2,u_5,u_6)$$

- Distance matrix D

	u_1	u_4	u_5	u_6
u_1	0	1	0.4	1.1
u_4	1	0	0.7	1.4
u_5	0.4	0.7	0	0.7
u_6	1.1	1.4	0.7	0

- Regional influence (for simplicity, $\gamma_R(\cdot) = 1$)

$$I_R(u_1) = 1 + 3/8 + 2/3 + 1/3 = 2.375$$
$$I_R(u_4) = 3/8 + 1 + 1/2 + 1/4 = 2.125$$
$$I_R(u_5) = 2/3 + 1/2 + 1 + 1/2 = 2.666$$
$$I_R(u_6) = 1/3 + 1/4 + 1/2 + 1 = 2.083$$

3. Computing Regional Influence

- Based on closeness centrality
- Set of edges weights W
  - For edge (u_x, u_y) of social graph G, $w_{xy} = -\ln p_{xy}$
  - Social distance $d(u_x, u_y)$, sum of weights on shortest path from u_x to u_y on G
- Propagation probability of an mip

$$p(\pi^*_{xy}) = e^{-d(u_x,u_y)}$$

- Regional influence

$$I_R(u_x) = \sum_{\forall u_i \in U_R} e^{-d(u_x,u_i)} \cdot \gamma_R(u_i)$$

4. The DRIC algorithm

Input: social graph $G(U,E)$; set of weights W; set of locations L; set of check-ins C; spatial region R; value k
Output: top-k list $\mathcal{T}$
Variables: set of regional users U_R, social distance matrix D

```
U_R ← GetRegionalUsers(U, L, C, R);
foreach u_i ∈ U_R do
    D ← Dijkstra(u_i, G, W, U_R);
    I_R(u_i) ← ComputeRegionalInfluence(u_i, U_R, D);
    push u_i to T;
return T;
```

5. Experiments

Dataset	$\|U\|$	$\|E\|$	$\|L\|$	$\|C\|$
Gowalla	197K	950K	1.3M	6.4M
Brightkite	58K	214K	773K	4.5M
Foursquare1	18K	116K	43K	2M
Foursquare2	11K	47K	187K	1.4M

- Response time (sec) varying query selectivity, $k = 5$

Dataset	$\|U_R\|/\|U\|$ (%)				
	0.1	0.2	0.3	0.4	0.5
Gowalla	140.6	262.1	432.5	590.6	1148.6
Brightkite	9.6	17.9	26.8	42.1	71.5
Foursquare1	0.9	2.3	3.2	5.8	11.2
Foursquare2	0.2	0.4	0.6	0.9	1.9