Participation Maximization Based on Social Influence in Online Discussion Forums

Tao Sun[1,2], Wei Chen [2], Zhenming Liu [2,3], YajunWang [2], Xiaorui Sun [2], Ming Zhang [1], Chin-Yew Lin [2]

[1]Peking University

[2]Microsoft Research Asia

[3]Harvard School of Engineering and Applied Sciences

Outline

- Motivation
- User Model based on social influence
- Participation Maximization
  - Problem formulation
  - Algorithms
- Summary

Motivation

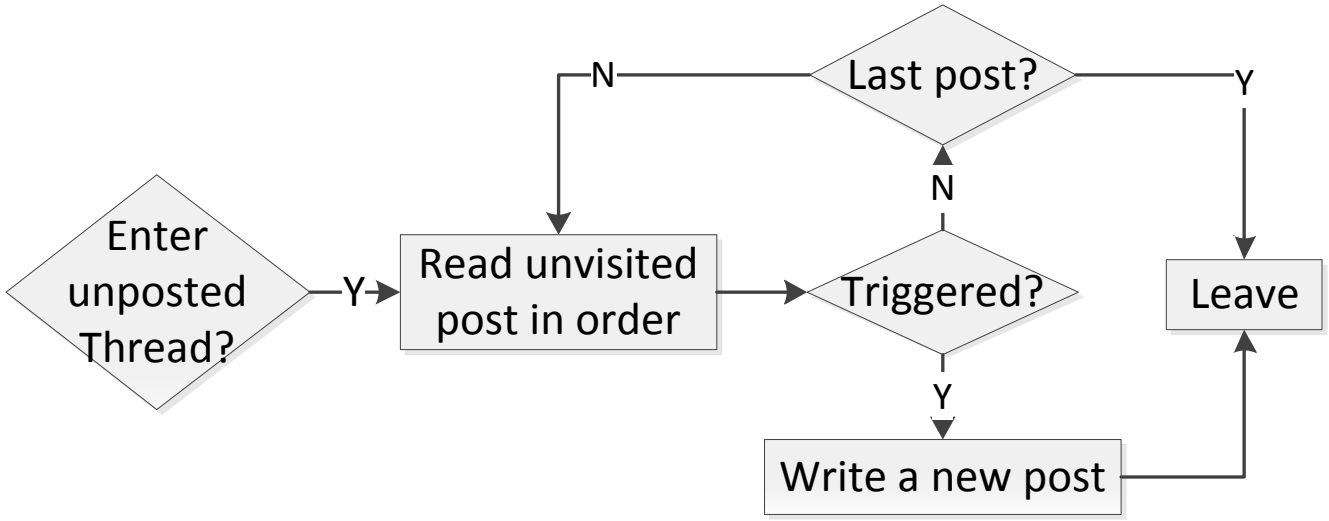
- Increase participation in online forums
  - Users expect a wide audience
  - Forum owners care about the traffic
- Users influence each other
  - Users tend to post after others
- Goal
  - Maximize participants based on social influence

Influence Network in forums

- Given an Influence network $G = (\mathcal{U}, E, w)$, directed, asymmetric and weighted
  - social ties, $u \rightarrow v$
  - the corresponding influence values, $W_{u,v}$
- Add a thread user τ to incorporate the influence from the original thread
  - $\tau \rightarrow v$: influence from the thread content
  - One τ for a set of threads on a specific topic, e.g., one category

User Model Based on Influence

- User posting model



Problem Description

Allocate B threads to each user, e.g., in his sidebar

↓

Users *visit* their suggested threads with a *higher probability*

↓

Users *post* in these threads with a *higher probability*

↓

Their posts *further influence* more users to post

↓

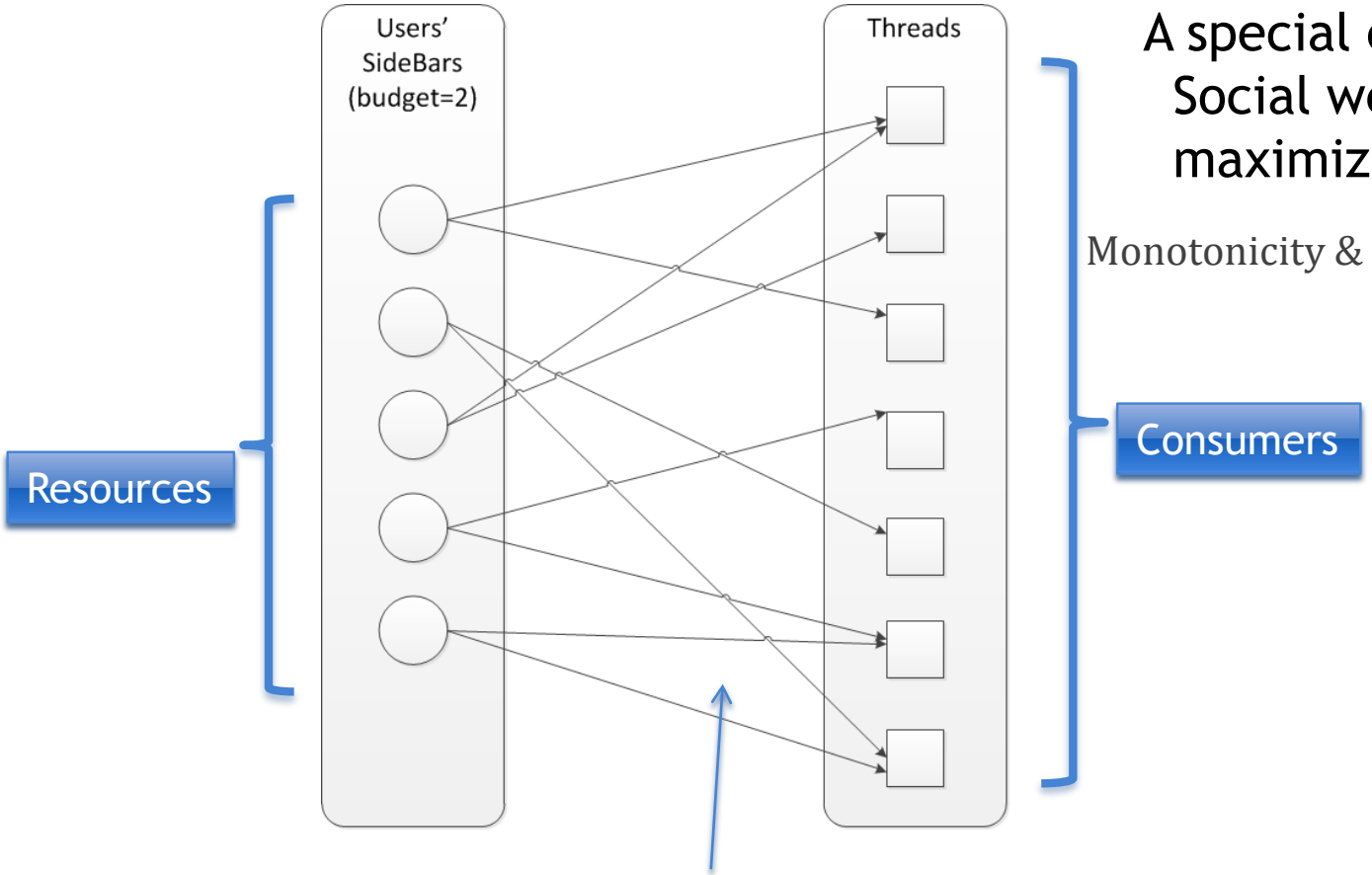
Increase the overall number of participants

How to allocate?

To maximize the total #participants through influence propagation.

Problem Formulation

- Participation Maximization



A special case of Social welfare maximization

Monotonicity & submodularity

Optimize the allocation to maximize social welfare (participants)

Comparison With Related Problems

Recommender Systems	Participation Maximization
Focus on users who are likely to post NOW	Consider FUTURE users who will be influenced to post

Influence Maximization (viral marketing)	Participation Maximization
Maximize participants in ONE specific threads	Maximize TOTAL participants in ALL threads

Thread Allocation Algorithms

- **Random allocation** – simple and fast
  - When all the consumers have *the same utility function*, achieve $\left(1-\frac{1}{e}\right)-approximation$
  - Works for *new threads*
    - Not for threads with participants already, different utility functions
- **Approximation algorithm** (RPA)
  - $\left(1-\frac{1}{e}\right)-approximation$, Dobzinski and Schapira 2006
  - Assume the computation of utility function is done by an oracle
  - Require $> (mn)^7$ iterations, **infeasible** for a real forum

Thread Allocation Algorithms

- Our heuristic algorithm:

  Thread Allocation based on Influence (TABI)

  - Symbols
    - EP_j - Existing Participants in T_j
    - I_v - the set of *v*'s *in-neighbors*
    - O_v - the set of *v*'s *out-neighbors*
    - δ_k - the original visit probability in kth time slot (discretize continuous time)
    - δ^*- the boosted probability that *v* visits threads in *v*'s SideBar
- v is influenced/activated in a suggested thread j
  - $\delta^* \left(1 - \prod_{u \in EP_j \cap I_v} (1 - w_{u,v})\right)$
- v influences others in *(k+1)th* time slot
  - $\delta_{k+1} \sum_{x \in O_v \setminus EP_j} w_{v,x} \prod_{u \in EP_j \cap I_x} (1 - w_{u,x})$

Thread Allocation Algorithms

- the *additional inf* Δinf_v^j that brought by displaying thread T_j to user *v*'s sidebar
  - (omitted visit pr which is same for all users)

$$\Delta inf_v^j = \underbrace{\left(1 - \prod_{u \in EP_j \cap I_v} (1 - w_{u,v})\right)}_{\text{v is activated}} \left(1 + \sum_{x \in O_v \setminus EP_j} w_{v,x} \prod_{u \in EP_j \cap I_x} (1 - w_{u,x})\right)$$

v itself

v's out-neighbors who are activated

Algorithm 2 TABI

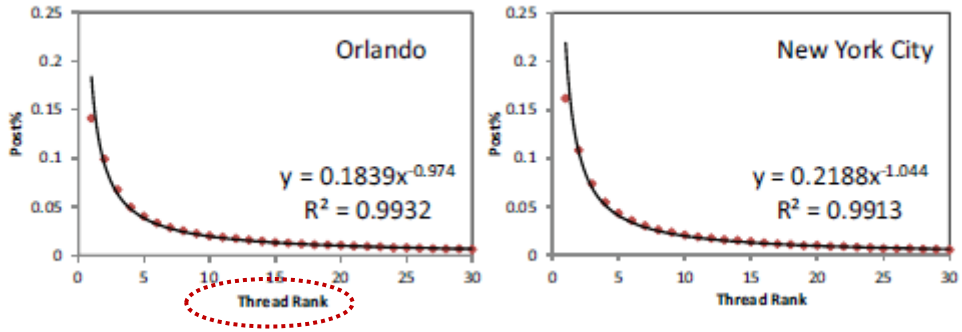
```
for each v ∈ U do
    for each j ∈ T do
        calculate ΔInf_v^j as Equation 2
    Rank threads by ΔInf_v^j in descending order
    Select top B threads to display in v's sidebar
```

Thread Allocation Algorithms

- **Personalized recommendation**
  - Topic sensitive early adoption based information flow --- **TEABIF**
  - Song, X., Tseng, B. L., et al. Personalized recommendation driven by information flow. In SIGIR '06.
- Target at different goals:
  - individual recommendation vs. overall participation
- Outcomes are comparable: **#participants**

Simulation based on User Model

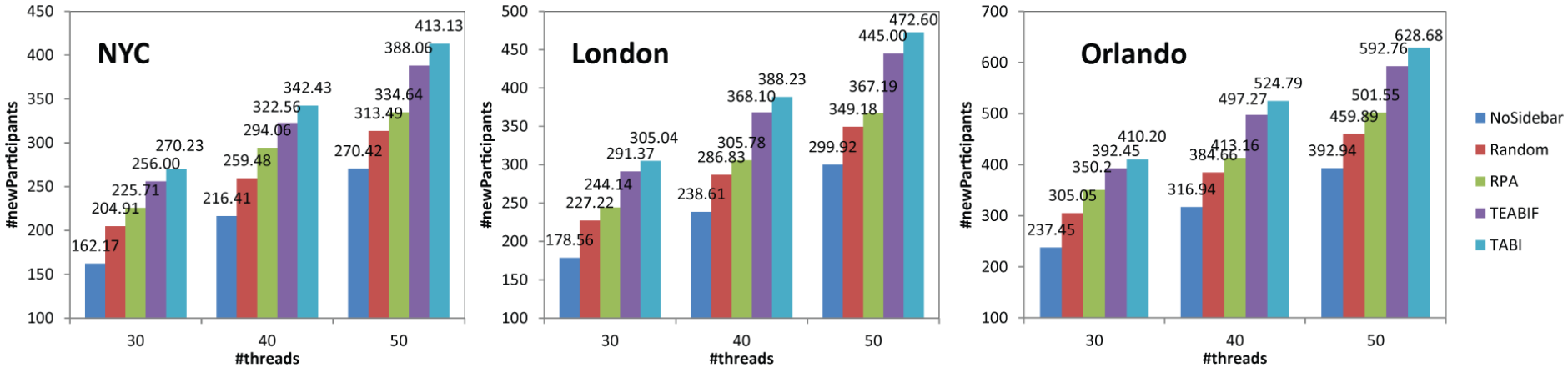
- Parameter Extraction (TripAdvisor)
- Implicit Influence Network
  - $G_\tau = (U_\tau, E_\tau, w)$
  - Keep edge u→v iff v follows u to post in at least N threads (N=2)
  - EM-algorithm to learn edge weight $w_{u,v}$ (Gruhl and Guha 2004)
- Visit Probabilities
  - Note Random, TEABIF and TABI don't rely on visit probabilities
  - δ_r, drawn from empirical distribution



Thread rank: thread T_j's rank in chronological order of all threads at a certain time

Different Allocation Algorithms

Intend to verify that TABI could perform consistently better under different numbers of threads (30, 40 and 50)

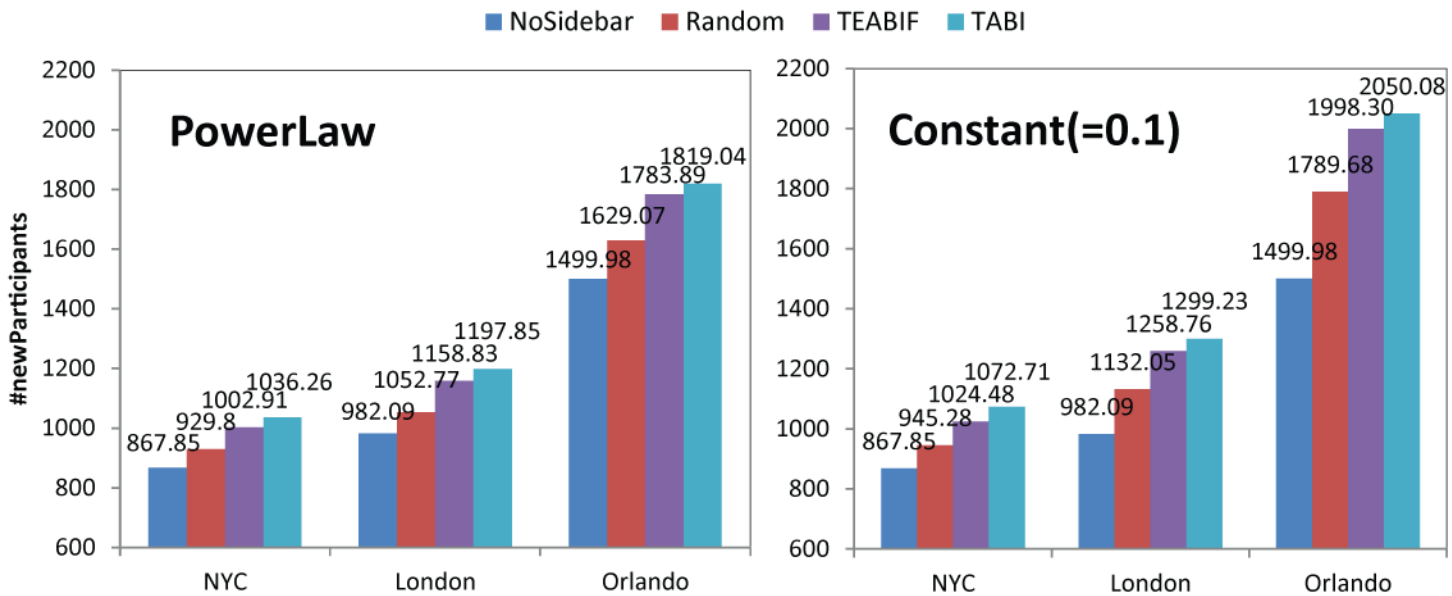


TABI outperforms all

Parameter setting:
B=5, allocation time slot s=2
δ_r to approximate δ_t
Boosted visit probability $\delta^* = 0.8$.
The value of δ^* won't affect thread allocation of Random, TEABIF and TABI
Simulation shows total participants under RPA is still linear to δ^*

Different Visit Probability

Intend to verify that TABI could perform consistently better under different original visit probabilities



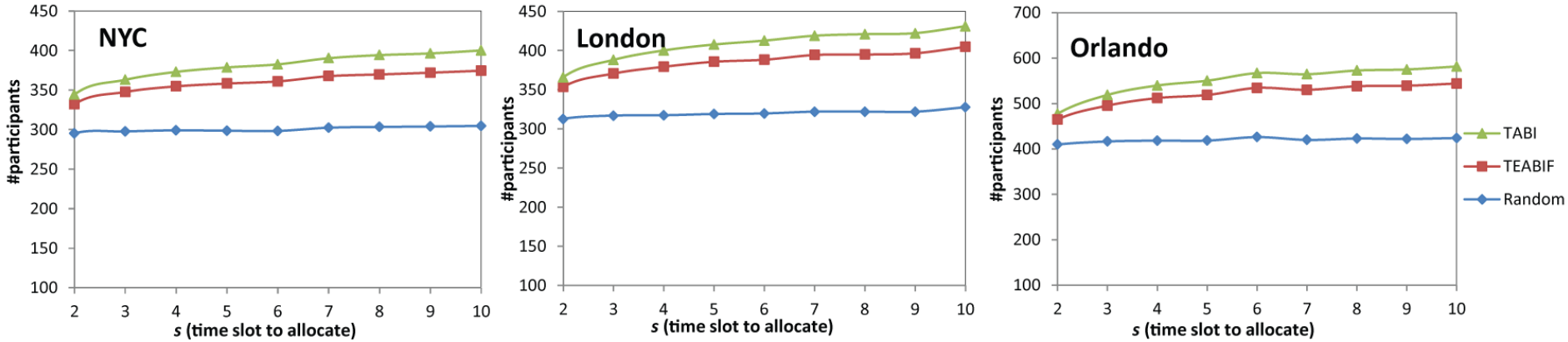
TABI outperforms all

Parameter settings:

1) Power Law: $\delta_t = kt^{-\alpha}$ $(k = 0.3,\ \alpha = 0.6)$,
   to simulate the decreasing trend with a larger visit probability compared to δ_r
2) Constant Value: $\delta_t = 0.1$ for all t

Different Allocation Time Slots

Intend to verify that TABI could perform consistently better when allocating in different time slots



TABI outperforms all

Seems that participation increases as s.
A larger boost in visit probabilities may provide more participation, but the selection of s should consider other factors such as user experience

Parameter settings:
Vary allocation time slot s from $s = 2$ to $s = 10$.

Other Applications

- Advertisements in Facebook

- Posts in Google Buzz
- Comments in YouTube
- Etc.

Summary

- Propose a personalized allocation mechanism to maximize total participation through influence propagation
  - Formulate the problem of participation maximization
    - Monotonicity and submodularity
  - Propose a heuristic algorithm--- TABI
    - Efficient, effective and robust

THANK YOU

Twitter: @GabriellaTaoSun