Learning user-specific latent influence and susceptibility from information cascades

Authors: Yongqing Wang, Huawei Shen, Shenghua Liu, Xueqi Cheng
Speaker: Yongqing Wang
Institute of Computing Technology, Chinese Academy of Sciences
In social media, users can receive news in time via such spontaneous information delivery way.
Cascade dynamics modeling

**Node:** user

**Edge:** propagation probability

### Basic Info.
- Structure of social network
- Record of information cascade
- Demographic/content characteristics of users

### Related works
- Holland, 71
- Granovetter, 73
- Barabasi, 99
- Kempe, 03
- Ugander, 13
- Leskovec, 07
- Saito, 08
- Goyal, 10
- Romero, 11
- Huang, 12
- Cui, 11
- Aral, 12
- Artzi, 12
- Liu, 12
- Tan, 14
The assumption of $n^2$ independent parameters makes pair-wise models suffering over-representation and over-fitting problems.

$p_{1,2}$, $p_{2,1}$ and $p_{2,3}$ are independent with each others
Our assumption makes user-specific model only need $2*n*d$ parameters which overcome flaws in pair-wise models.

$$I_1, S_1 \quad I_2, S_2 \quad \text{inferred} \quad S_3$$

**LIS model**

**Sender**

$I$: influence

**Receiver**

$S$: susceptibility

$$I^T S$$

Interpersonal influence

$$1 - \exp\left(-\lambda I^T S\right)$$

Propagation probability
Basic concepts

Notations:

- Message $m$
- Cascade $C^m : (a_1^m, \cdots, a_N^m)$

\[ (u_3, u_1, u_2, u_4, u_5) \] Ranked by ascending order of activation time

Basic rules:

- One user can try to activate others only once
- One user can be activated only once

Cascade context:

- User’s activating attempt depends on historical influencers
LIS model

Message $m$

cascade: $u_3 \xrightarrow{\delta} u_2 \xrightarrow{\delta} u_4 \xrightarrow{\delta} u_5$

cascade context: $\{u_3\}, \{u_3, u_2\}$

timeline

1. When one user is activated, he has one chance to activate its direct neighbors.

   $\delta(u,v) = \begin{cases} 
   1, & \text{if} \ (u,v) \text{ has an directed edge} \\
   0, & \text{if} \ (u,v) \text{ has no relationship} 
   \end{cases}$

2. Whether his attempt succeeds depends on the cascade context at that time.

   Cascade context: $D_{v,i}^m = \{a_j^m \mid j \leq i, \delta(a_j^m, v) = 1\}$

   p.s. the length of cascade context is controllable.

Likelihood of $u_4$'s status chain:

\[
P(z_v^m \mid \delta) = p(z_{v,0}^m) \prod_{i=1}^{N} p(z_{v,i}^m \mid z_{v,i-1}^m, D_{v,i}^m, \delta)
\]

\[
p(z_{v,0}^m = 1) = \begin{cases} 
   1, & \text{if} \ v \text{ is the source} \\
   0, & \text{otherwise}
   \end{cases}
\]

\[
p(z_{v,i}^m = 1 \mid z_{v,i-1}^m = 0, D_{v,i}^m, \delta) = 1 - \exp \left\{ -\lambda \delta(a_i^m, v) \sum_{u \in D_{v,i}^m} I_u^T S_v \right\}
\]
Graphical model & optimization

Graphical representation of LIS model for one node

Parameter estimation

**Input**: Collection of cascades observed in a given time period

**Output**: User-specific influence and susceptibility $I$, $S$

Construct diffusion network $\delta$ from cascades

Initialize parameters with random values, including $I$, $S$

Repeat

for $i=1$ to $n$

Calculate gradient $\frac{\partial \mathcal{L}}{\partial I_u}$ and $\frac{\partial \mathcal{L}}{\partial S_v}$

end for

Update $I$ and $S$ with PG method

Until maximum epoch $M$ is reached or gradient vanish

Objective function:

$$
\mathcal{L}(C) = - \sum_{v \in V} \sum_{D_{v,i} \in \mathcal{P}(v)} \left( n_{z_{v,i}, D_{v,i}} \log p(z_{v,i} | z_{v,i-1}, D_{v,i}, \delta) \right)
+ \gamma_I \left\| I \right\|_F^2 + \gamma_S \left\| S \right\|_F^2
$$

s.t. $I_{ij} \geq 0, S_{ij} \geq 0, \forall i, j$
Datasets

Synthetic data

1) BA network, #nodes=1000;
2) The shuffle network

Networks

$I, S : f(x) = \frac{1}{2} \sqrt{x}, x \sim U(0, 1)^5$

Parameters

Node

Setups

20%

80%

- training data
- test data
Datasets

Real data (Sina Weibo)

NETWORKS

DATA STATISTICS

|                | Training data                  | Test data               |
|----------------|--------------------------------|-------------------------|
|                | cascades                       | period                  |
| D1             | 395,852                        | 01/01~01/15             |
| D2             | 453,356                        | 01/16~01/31             |
| D3             | 386,152                        | 02/01~02/15             |
| T1             | 160,868                        | 01/16~01/20             |
| T2             | 122,509                        | 02/01~02/05             |
| T3             | 145,143                        | 02/16~02/20             |

Aggregate from all cascade graphs
Experimental setups

Baselines

• Expectation Maximization estimation (EM)
• Static Bernoulli model (SB)
• Static Jaccard model (SJ)

Prediction tasks

• Cascade dynamics prediction
• Cascade size prediction
• “who will be retweeted” prediction

post-process by: matrix factorization method
Cascade dynamics prediction

Cascade dynamics prediction directly reflect models’ abilities on describing information cascade.

**Synthetic data**

|                  | UB | LIS | SB | SJ | EM |
|------------------|----|-----|----|----|----|
| BA network       | 0.659 | 0.654 | 0.607 | 0.618 | 0.561 |
| The shuffle one  | 0.659 | 0.608 | 0.509 | 0.525 | 0.507 |

The AUC resulted by the LIS model is closer to UB, and the LIS model is more stable than pairwise models.

*p.s. UB refers to upper bound*

**Real data (Sina Weibo)**

The AUCs decrease dramatically on pairwise models, when they suffer over-fitting problem.

The LIS model performs better as the increase of the length of cascade context. The AUCs decrease dramatically on pairwise models, when they suffer over-fitting problem. The AUC resulted by the LIS model is closer to UB, and the LIS model is more stable than pairwise models.

Cascade dynamics prediction directly reflect models’ abilities on describing information cascade.
Cascade size prediction is one of the most important applications for cascade dynamics modeling.

|       | LIS ($l=0$) | LIS ($l=3$) | LIS ($l=5$) |
|-------|-------------|-------------|-------------|
| T1    | 0.163 ± 0.0133 | 0.140 ± 0.0155 | 0.141 ± 0.0217 |
| T2    | 0.287 ± 0.0093 | 0.280 ± 0.0080 | 0.286 ± 0.0065 |
| T3    | 0.095 ± 0.0150 | 0.094 ± 0.0150 | 0.097 ± 0.0093 |

|       | SB          | SJ          | EM          |
|-------|-------------|-------------|-------------|
| T1    | 0.191 ± 0.0190 | 0.524 ± 0.0046 | 0.258 ± 0.0160 |
| T2    | 0.333 ± 0.0099 | 0.621 ± 0.0048 | 0.338 ± 0.0387 |
| T3    | 0.171 ± 0.0388 | 0.505 ± 0.0450 | 0.189 ± 0.0112 |
Prediction of “who will be retweeted”

The prediction of “who will be retweeted” is one way to examine interpersonal influence under quantitative understanding.

|                | LIS (l=5) | SB  | SJ  | EM  |
|----------------|-----------|-----|-----|-----|
| **Acc(%)**     |           |     |     |     |
| T1             | 58.48     | 57.02| 49.99| 53.48|
| T2             | 57.61     | 55.05| 49.65| 52.23|
| T3             | 59.58     | 55.38| 50.85| 55.41|
| **MRR**        |           |     |     |     |
| T1             | 0.791     | 0.784| 0.748| 0.766|
| T2             | 0.786     | 0.773| 0.745| 0.758|
| T3             | 0.797     | 0.775| 0.752| 0.775|

“who will be retweeted” list

Interpersonal influence

The predicted one

other users
Summary

- Propose **LIS model** to depict cascade dynamics
  - Model user-specific latent influence and susceptibility
- Overcome *over-representation* and *over-fitting* problems in pair-wise models
- Capture **context-dependent factors** like cumulative effects in information propagation
- Design effective algorithm to train the model and well apply to key prediction tasks on information propagation
Authors

Yongqing Wang  Huawei Shen  Shenghua Liu  Xueqi Cheng
Network Analysis and Social Computing (NASC) group

http://www.nascgroup.org/members

Yongqing Wang
Huawei Shen
Wei Chen
Junming Huang
Xiaqian Sun

Peng Bao
Suqi Cheng
Tong Man
Jianye Yu
Bingjie Sun
Thanks 😊

Homepage: http://yongqwang.com
or http://114.215.84.52/
Email: wangyongqing@software.ict.ac.cn