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

<table>
<thead>
<tr>
<th>Related works</th>
<th>Basic Info.</th>
<th>Propagation probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holland,71</td>
<td>Structure of social network</td>
<td>Record of information cascade</td>
</tr>
<tr>
<td>Granovetter,73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barabasi,99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kempe,03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kempe,03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ugander,13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Related works</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leskovec,07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saito,08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goyal,10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romero,11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huang,12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu,12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cui,11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aral,12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artzi,12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tan,14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pairwise models

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 other.

Inferred: $p_{1,3} = 0$
User-specific model

Our assumption makes user-specific model only need $2 \times n \times d$ parameters which overcome flaws in pair-wise models.

$$I_1, S_1 \quad I_2, S_2 \quad S_3$$

inferred

LIS model

Sender

$I$: influence

Receiver

$S$: susceptibility

Interpersonal influence

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

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: $u_3 \times u_2 u_4 u_5$

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

timeline

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

\[
\delta(u, v) = \begin{cases} 
1, & \langle u, v \rangle \text{ has a directed edge} \\
0, & \langle u, v \rangle \text{ has no relationship}
\end{cases}
\]

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

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

_p.s. the length of cascade context is controllable._

Likelihood of $u_4$’s status chain:

\[
P(z_v^m \mid \delta) = p(z_v^m_0) \prod_{i=1}^{N} p(z_v^m_i \mid z_v^m_{i-1}, D_{v,i}^m, \delta)
\]

\[
p(z_v^m_0 = 1) = \begin{cases} 
1, & v \text{ is the source} \\
0, & \text{otherwise}
\end{cases}
\]

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

Graphical representation of LIS model for one node

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 \( n \) do
    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:
\[
\mathcal{L}(C) = -\sum_{v \in V} \sum_{D_{v,i} \in \mathcal{D}} \left( n_{v,i} \cdot p_{v,i} \log p\left(z_{v,i} \mid z_{v,i-1}, D_{v,i}, \delta\right) \right)
+ \gamma_I \| I \|_F^2 + \gamma_S \| S \|_F^2
\]
subject to \( I_{ij} \geq 0, S_{ij} \geq 0, \forall i, j \)
Datasets

**Synthetic data**

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

**NETWORKS**

**PARAMETERS**

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

**SETUPS**

- 20% training data
- 80% test data
Datasets

Real data (Sina Weibo)

NETWORKS

DATA STATISTICS

<table>
<thead>
<tr>
<th></th>
<th>Training data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cascades</td>
</tr>
<tr>
<td>D1</td>
<td>395,852</td>
</tr>
<tr>
<td>D2</td>
<td>453,356</td>
</tr>
<tr>
<td>D3</td>
<td>386,152</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cascades</td>
</tr>
<tr>
<td>T1</td>
<td>160,868</td>
</tr>
<tr>
<td>T2</td>
<td>122,509</td>
</tr>
<tr>
<td>T3</td>
<td>145,143</td>
</tr>
</tbody>
</table>
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
Cascade dynamics prediction

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

**Synthetic data**

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

<table>
<thead>
<tr>
<th></th>
<th>UB</th>
<th>LIS</th>
<th>SB</th>
<th>SJ</th>
<th>EM</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA network</td>
<td>0.659</td>
<td>0.654</td>
<td>0.607</td>
<td>0.618</td>
<td>0.561</td>
</tr>
<tr>
<td>The shuffle one</td>
<td>0.659</td>
<td>0.608</td>
<td>0.509</td>
<td>0.525</td>
<td>0.507</td>
</tr>
</tbody>
</table>

**Real data (Sina Weibo)**

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.

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

p.s. UB refers to upper bound
Cascade size prediction

Cascade size prediction is one of most important applications for cascade dynamics modeling.

### MAPE (mean absolute percentage error) values

<table>
<thead>
<tr>
<th></th>
<th>LIS ((l=0))</th>
<th>LIS ((l=3))</th>
<th>LIS ((l=5))</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>0.163 ± 0.0133</td>
<td>0.140 ± 0.0155</td>
<td>0.141 ± 0.0217</td>
</tr>
<tr>
<td>T2</td>
<td>0.287 ± 0.0093</td>
<td>0.280 ± 0.0080</td>
<td>0.286 ± 0.0065</td>
</tr>
<tr>
<td>T3</td>
<td>0.095 ± 0.0150</td>
<td>0.094 ± 0.0150</td>
<td>0.097 ± 0.0093</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SB</th>
<th>SJ</th>
<th>EM</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>0.191 ± 0.0190</td>
<td>0.524 ± 0.0046</td>
<td>0.258 ± 0.0160</td>
</tr>
<tr>
<td>T2</td>
<td>0.333 ± 0.0099</td>
<td>0.621 ± 0.0048</td>
<td>0.338 ± 0.0387</td>
</tr>
<tr>
<td>T3</td>
<td>0.171 ± 0.0388</td>
<td>0.505 ± 0.0450</td>
<td>0.189 ± 0.0112</td>
</tr>
</tbody>
</table>
The prediction of “who will be retweeted” is one way to examine interpersonal influence under quantitative understanding.

<table>
<thead>
<tr>
<th>Acc(%)</th>
<th>LIS ($l=5$)</th>
<th>SB</th>
<th>SJ</th>
<th>EM</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>58.48</td>
<td>57.02</td>
<td>49.99</td>
<td>53.48</td>
</tr>
<tr>
<td>T2</td>
<td>57.61</td>
<td>55.05</td>
<td>49.65</td>
<td>52.23</td>
</tr>
<tr>
<td>T3</td>
<td>59.58</td>
<td>55.38</td>
<td>50.85</td>
<td>55.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MRR</th>
<th>LIS ($l=5$)</th>
<th>SB</th>
<th>SJ</th>
<th>EM</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>0.791</td>
<td>0.784</td>
<td>0.748</td>
<td>0.766</td>
</tr>
<tr>
<td>T2</td>
<td>0.786</td>
<td>0.773</td>
<td>0.745</td>
<td>0.758</td>
</tr>
<tr>
<td>T3</td>
<td>0.797</td>
<td>0.775</td>
<td>0.752</td>
<td>0.775</td>
</tr>
</tbody>
</table>
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
Thanks 😊

Homepage: [http://yongqwang.com](http://yongqwang.com)
or [http://114.215.84.52/](http://114.215.84.52/)

Email: wangyongqing@software.ict.ac.cn