

Quantifying Structural Patterns of Information Cascades



Chengxi Zang, Peng Cui, Chaoming Song, Christos Faloutsos, and Wenwu Zhu zangcx13@mails.tsinghua.edu.cn, chaomingsong@gmail.com, christos@cs.cmu.edu, {cuip, wwzhu}@tsinghua.edu.cn

Real Information Cascades exhibit rich structures



Dataset

Tencent Weibo (432 million cascades, 101 million users, 7day-period, full scale, explicit spreading traces)

Cascade Structure Definition

- The structure of a cascade C is the directed graph (V, E):
- **Node:** $u \in V$ is a retweeter, u_0 denotes the original poster.
- **Edge:** $(u, v) \in E$, user v retweets user u.
- Weight of edge (u, v): w(u, v), v retweets u w(u, v) times. **Self-Loop** edge (u, u): u retweets himself.

Findings:

• Structural complexity of information cascades is far beyond the previous conjectures: e.g., Star-like, narrow-and-deep-like, etc.

What do Information Cascades look like in the Size Metric Space

Reciprocal edge $r_{u,v} = \{(u,v), (v,u) \mid u \neq v\} u$ and v retweeted each other.

Size metrics







Orientation metrics

- **Branch coefficient:** $\frac{std(k_{out})}{dt}$ $< k_{out} >$ **Converge coefficient:** $\frac{std(k_{in})}{d}$ $\langle k_{in} \rangle$ Self-loop ratio: $\frac{|\{u|(u,u)\in E\}|}{|||}$
 - **Reverse ratio:** $\frac{|\{(u,v) \mid (u,v) \in E \& (v,u) \in E \& u \neq v\}|}{|\{(u,v) \mid (u,v) \in E \& (v,u) \in E \& u \neq v\}|}$

What do Information Cascades look like in the Orientation Metric Space

- Fat-tailed nature
 - Max (mass, breadth, length) = (1414815, 1408024, 57).
- Breadth accounts for a large portion of mass.
- Large cascades are either deep or wide.
- Majority of cascades are wide and shallow.
- There exist narrow and deep cascades.
- No wide and deep, or very narrow and deep cascades.





- The converging-in, reverse and self-loop orientations are ubiquitous besides the branching-out orientation.
- Bimodal distribution for branching-out orientation.
- Down-stair-distribution for converging-in, reverse and self-loop orientations.
- Extreme cases exist: each node has a self-loop edge, each edge has a reverse edge.
- Non-coexistence relationship between branching-out and converging-in orientations.



