Enhancing Structural Diversity in Social Networks by Recommending Weak Ties

**Motivation**

- **Beyond accuracy**
  - Novelty & diversity
  - Many notions from social network analysis
  - Structural diversity → weak ties

**Structural diversity**

- **Community edge Cini complement (CEGC)**
  - Consider redundancy between weak ties
  - Analysis of links existing communities
  - Low CEGC → Skewed distribution → Low diversity

- **Local redundancy**: Transitive closure
  - Triadic closure: smallest unit of structural redundancy
  - Clustering coefficient complement (CECC)

**Effect on information diffusion**

- **Hypothesis**
  - The more structurally diverse is the recommendation, the more diverse and novel (non-redundant) will be the information flow through the network

**Experiment description**

- **Start with a well-behaved baseline** → Implicit MF (most accurate method)
- **Berkank baseline** to enhance a structural metric of the network
- **Simulate the flow of information through the extended network**
- **Analyze properties of diffusion (speed, novelty & diversity)**

**Data**

- Same networks as the ones used for the recommendation experiments
- **Information to propagate**: Tweets
  - originally published after the temporal split
  - containing hashtags which appear in at least 10 different tweets (avoid noise)

**Protocol**

- Information is propagated to all followers
- **User** retweets a tweet only if she retweeted it in real life → deterministic

**Metrics enhancement**

- **Enhance a global property μ of the network**
- **Berkank baseline recommendation by greedy maximization of objective function**

\[ \phi(f, i, j, \mu) = (1 - \lambda) \sum_{(v, \mu_i)} f(v, \mu_i) + \mu \phi(i, j) \]

**Algorithm: Global greedy reranking**

1. **Input**
   - \( E \) \& \( G \) \& \( \phi(i, j) \)'s
2. **Output**
   - \( \phi_{opt}(E, G, \phi(i, j), \mu) \)
3. **Algorithm**
   - **Initialize**
     - \( \phi = \phi_{opt}(E, G, \phi(i, j), \mu) \)
   - **Iterate**
     - Until no more improving recommendation
     - **For each**
       - \( i \) in \( G \)
     - **Do**
       - **Initialize**
         - \( \phi' = \phi_{opt}(E, G, \phi(i, j), \mu) \)
       - **For each**
         - \( j \) in \( G \)
       - **Do**
         - **If**
           - \( \phi(f(i, j)) > \phi(f(i, k)) \)
         - **Then**
           - \( \phi = \phi' \)
   - **Return**

**Recommendation experiments**

- **How do state of the art algorithms perform in terms of structural diversity?**

<table>
<thead>
<tr>
<th>Recommendation algorithms</th>
<th>Original</th>
<th>( \phi(E) )</th>
<th>( \phi(G) )</th>
<th>( \phi(E, G) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implicit MF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personalized SALSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive-Adapted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaccard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Popularity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralized CB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baselines: random, popularity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Information diffusion properties**

- **Notation**
  - \( I_C \) Set of all hashtags
  - \( t \) a tweet defined as a subset of \( I_C \)
  - \( a_t \) at time \( t \)
  - Has received the tweets \( t \) containing the hashtags \( N_{t}(t) \)
  - \( |N_{t}(t)| \)

- **Speed**
  - Most analyzed network efficiency feature in diffusion processes
  - How many tweets are propagated and received?

- **Novelty and diversity**
  - Measured in terms of hashtags
  - **Novelty**
    - How novel is the information received by users?
    - Internal ranking factor (IRF)
  - **Diversity**
    - Are hashtags evenly distributed over the population?
    - Potential to incentivize Ellen hubs
    - Ranking time complement (RSC)

- **Recommendation results**
  - IECG provides the best trade-off between accuracy, structural properties and information diversity
  - Recommending weak ties improves the novelty of the information received by the different users