Measuring Friendship Closeness: A Perspective of Social Identity Theory

Shiqi Zhang, Jiachen Sun, Wenqing Lin, Xiaokui Xiao, Bo Tang

October 2022

CIKM 2022
Outline

• Problem and Applications
• Existing Works
• Proposed Measures
• Experiments
• Deployments
Problem Definition

• Given a graph $G = (V, E)$
  - $v_s \in V$: the user in the social network
  - $(v_s, v_t) \in E$: two users are friends in the social network

• We aim to measure $\text{TFC}$ (Topological Friendship Closeness) for each friend pair $(v_s, v_t) \in E$. 
Application Scenario

- Friendship-enhancing event in Tencent’s Games
  - source $v_s$: the user who sent the invitation
  - target $v_t$: the user who received the invitation

- TFC is applied for (i) user behavior understanding and (ii) target recommendation.
Existing Solutions

a. Individual-level measures
   - Tie strength, # common friend
   - Personalized PageRank, similarity between node embeddings

b. Group-level measures
   - Structural diversity, user-group tie strength

Our objective: explore more in-group information
Proposed Measures: Theory support

- Social identity theory (SIT)
  - The inclination that a target endorses behaviors of users inside the same group is affected by psychological factors.

| Factor            | Meaning                                                                 |
|-------------------|-------------------------------------------------------------------------|
| Multi-membership  | Number of groups                                                        |
| Inclusiveness     | Number of in-group members                                               |
| Solidarity        | $v_t$’s psychological bond with in-group members                         |
| Centrality        | Importance of a group in $v_t$’s cognition                               |
| Self-stereotyping | Similarity of $v_t$ and group average in $v_t$’s cognition               |
| In-group homogeneity | Similarity within a group                                                |
| Social standing   | Social standing of a group                                               |

- Novelty: import psychological factors for TFC measures
- Challenge: how to reflect these factors by structural information?
Proposed Measures: Overview

- Social categorization
  - Candidate group = CC in the ego network of the target.

- **SIT-based measure definition**
  - Define quantitative measures to describe each factor.

- Inclination inference
  - Supervised learning via XGBoost.
Proposed Measures: Definitions

- \((C):\) the group that the given source and target are in
- Abstract the edge as a spring
  - tie strength \(w\) as the stiffness constant.
  - similarity \(\delta\) as the displacement.
- **UGT** (solidarity, self-stereotyping)
  - The average attractive force from users in \(C\) to the target.
- **IGT** (ingroup-homogeneity)
  - The average attractive force among users in \(C\).
Experiments: Datasets and Setup

Table 2: Dataset statistics ($M = 10^6$, $B = 10^9$).

| Dataset | $|V|$  | $|E|$  | $|S|$  | $|T|$  |
|---------|------|------|------|------|
| FPS     | 77.2M| 1.1B | 33.5M| 43.6M|
| MOBA-A  | 111.0M| 4.5B | 111.0M| 94.7M|
| MOBA-B  | 130.2M| 6.5B | 120.5M| 99.7M|

• Individual-level competitors:
  – tie strength (Tie); #common friend (COM); Personalized PageRank (PPR); similarity between Node2vec embeddings (N2V)

• Group-level competitors:
  – Structural diversity (#CC); user-group tie strength (GT); in-group edge density (GD)

• UGT,IGT: $w$: Tie; $\delta$: N2V.
Experiments: Behavior Prediction

| Measure | Adoption | | | Invitation | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FPS | MOBA-A | | FPS | MOBA-A | | FPS | MOBA-A | | FPS | MOBA-A | |
| | AUC | Accuracy | F1 score | AUC | Accuracy | F1 score | AUC | Accuracy | F1 score | AUC | Accuracy | F1 score |
| Tie | 0.7154 | 0.6965 | 0.6554 | 0.6017 | 0.6021 | 0.3958 | 0.6072 | 0.5985 | 0.4607 | 0.5361 | 0.5353 | 0.2200 |
| COM | 0.5488 | 0.5538 | 0.5615 | 0.5667 | 0.5576 | 0.6219 | 0.5456 | 0.5323 | 0.4674 | 0.5332 | 0.5281 | 0.5565 |
| PPR | 0.6565 | 0.6036 | 0.5596 | 0.5562 | 0.5388 | 0.4447 | 0.6289 | 0.5972 | 0.5786 | 0.5846 | 0.5589 | 0.5467 |
| N2V(cos) | 0.6976 | 0.6610 | 0.7171 | 0.5808 | 0.5626 | 0.5420 | 0.5608 | 0.5537 | 0.5683 | 0.5770 | 0.5630 | 0.5426 |
| N2V(euc) | 0.7076 | 0.6652 | 0.7091 | 0.5664 | 0.5566 | 0.5390 | 0.5679 | 0.5588 | 0.5628 | 0.5739 | 0.5585 | 0.5375 |
| #CC | 0.6153 | 0.5897 | 0.5378 | 0.5452 | 0.5288 | 0.4136 | 0.6091 | 0.5820 | 0.5392 | 0.5790 | 0.5551 | 0.5662 |
| GT | 0.6985 | 0.6572 | 0.6004 | 0.6295 | 0.5959 | 0.5777 | 0.5738 | 0.5652 | 0.3988 | 0.5397 | 0.5297 | 0.4875 |
| GD | 0.6077 | 0.5736 | 0.5269 | 0.6039 | 0.5728 | 0.5908 | 0.5811 | 0.5507 | 0.4490 | 0.5674 | 0.5508 | 0.4991 |
| SIT | 0.7995 | 0.7206 | 0.7350 | 0.7410 | 0.6780 | 0.6638 | 0.7307 | 0.6719 | 0.6754 | 0.6550 | 0.6086 | 0.6047 |

- Predictions for target adoption and source invitation
- **SIT outperforms all competitors** on two datasets in terms of two prediction tasks and three evaluation metrics.
Experiments: Conversion analysis

- Conversion rate: fraction of inviters/adopters in given rank.
- The conversion is more sensitive to #CC, GPR, GPPR, UGT

**Figure 4:** Conversion probability of adoption behaviors conditioned on each SIT-based measure in FPS.

**Figure 5:** Conversion probability of invitation behaviors conditioned on each SIT-based measure in FPS.
Deployments: Target Recommendation

Table 7: Online performance in MOBA-A.

| Measure   | Tie   | COM  | PPR   | N2V(euc) | SIT   |
|-----------|-------|------|-------|----------|-------|
| E2E rate  | 0.1018| 0.0958| 0.1066| 0.0739   | 0.1431|

Table 8: Online performance in MOBA-B.

| Measure   | Tie   | PPR   | SIT   |
|-----------|-------|-------|-------|
| E2E rate  | 0.1152| 0.1218| 0.1384|

- E2E rate: **target friends adopting the invitations**
  **source users seeing the event**.
- SIT improves the best competitor by up to **34.2%**.
Conclusion

• We propose six new TFC measures based on the social identity theory.

• The SIT-based measures are sensitive to user behavior conversion.

• The SIT-based measures can experimentally outperform the competitors.

• The SIT-based measures have been deployed to more than 10 friendship-enhancing events up till now.
Thank You!