Ranking Users by their Influence in Social Networks

Nouamane Arhachoui

Sorbonne Université, CNRS – LIP6, Paris, France

May 25th, 2023
Lake Como School of Advanced Studies, Como, Italy

Outline

Node importance

Various applications of Network Science require measuring the importance of a node:

- communication networks,
- social networks,
- urban networks,
- and more.

Example

Example

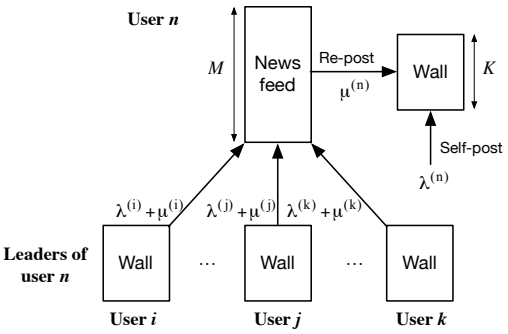
Most common centrality metrics may not be the best choice for networks with node interaction over time.

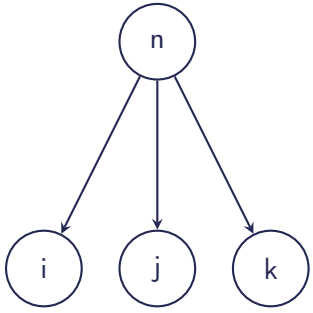
"Are social links valid indicators of real user interaction?", C. Wilson et al. (2012)

$\Longrightarrow$ it is necessary to go beyond pure structural metrics

Given a network where nodes create and share information over time, how can we measure the influence of a given node over the entire network?

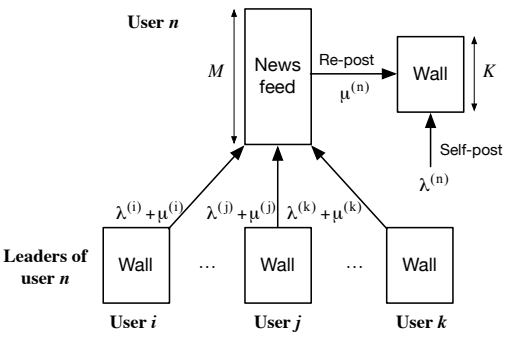
Social Platform Model

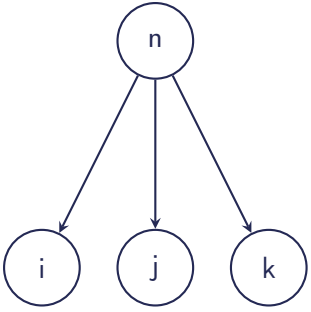




- $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $(i, j) \in \mathcal{E}$ iff user i follows user j. $|\mathcal{N}| = N$.
- Each user has a set of **followers** $\mathcal{F}^{(n)}$ and a set of **leaders** $\mathcal{L}^{(n)}$.
- Each user has 2 queues: a **Wall** of size K and a **Newsfeed** of size M.
- Each user has a **posting rate** $\lambda^{(n)}$ (posts created by n per unit of time) and a **re-posting rate** $\mu^{(n)}$ (posts that n shares per unit of time).

Social Platform Model



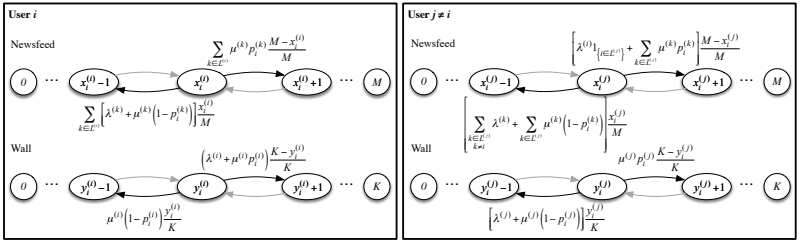


- Rate of post arrival on n's Wall: $\lambda^{(n)} + \mu^{(n)}$
- Rate of post arrival on n's Newsfeed: $\sum_{l \in \mathcal{L}^{(n)}} (\lambda^{(l)} + \mu^{(l)})$

Focus on posts of origin i

By focusing on post created by user i, we build two continuous Markov chains with the states:

- $x_i^{(n)}$: the number of posts of origin i that appear on the **Newsfeed** of n
- $y_i^{(n)}$: the number of posts of origin i that appear on the **Wall** of n



The ψ-score

Influence of user i

The steady state of this Markovian model leads to:

$$\psi_i = \frac{1}{N} \sum_{n=1}^{N} q_i^{(n)}$$

where,

- ψ_i: the ψ-score (influence of user i over the entire network)
- $q_i^{(n)}$: the expected number of posts created by i that are on the Newsfeed of n

Relation with PageRank

SORBONNE UNIVERSITÉ

Theorem

When all the users have the same activity, i.e. $\forall n \in [\![1, N]\!]\ \lambda^{(n)} = \lambda$ and $\mu^{(n)} = \mu$ and if $\frac{\mu}{\lambda+\mu} = \alpha \in [0, 1]$, then ψ-score = PageRank with a damping factor α

$\Longrightarrow$ The ψ-score is a **generalization of the PageRank** that uses the activity of nodes as additional information to measure the influence.

The psi-$score$ Python package

- An open-source project*: https://github.com/NouamaneA/psi-score
- Easy to use with a ScikitLearn-like API

```
>>> from psi_score import PsiScore
>>> adjacency = {0: [1, 3], 1: [0, 2], 2: [0, 1, 3], 3: [0]}
>>> lambdas = [0.23, 0.50, 0.86, 0.19]
>>> mus = [0.42, 0.17, 0.10, 0.37]
>>> psiscore = PsiScore()
>>> scores = psiscore.fit_transform(adjacency, lambdas, mus)
>>> scores
array([0.21158803, 0.35253745, 0.28798439, 0.14789014])
>>> np.round(scores, 2)
array([0.21, 0.35, 0.29, 0.15])
```

*Any contribution to the project is welcome!

Conclusion and Future work

The ψ-score

- allows ranking users by their influence
- generalizes the PageRank when user activity is heterogeneous

Future work:

- Explore generalizations of the ψ-score in time evolving networks
- Study the effect of day/night (or ON/OFF) user activity on the model
- Use the OSP model and the ψ-score to detect malicious behaviors

References


Arhachoui, Nouamane et al. (2022). "A Fast Algorithm for Ranking Users by their Influence in Online Social Platforms". In: *2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 526–533. doi: 10.1109/ASONAM55673.2022.10068673.


Giovanidis, Anastasios et al. (2021). "Ranking Online Social Users by Their Influence". In: *IEEE/ACM Transactions on Networking* 29.5, pp. 2198–2214. doi: 10.1109/TNET.2021.3085201.