Emergence of communities in weighted social networks

J. M. Kumpula\textsuperscript{1}, J.-P. Onnela\textsuperscript{2,1},
J. Saramäki\textsuperscript{1}, K. Kaski\textsuperscript{1} and J. Kertész\textsuperscript{3,1}

\textsuperscript{1}Laboratory of Computational Engineering, Helsinki University of Technology
\textsuperscript{2}Department of Physics & Saïd Business School, Oxford University
\textsuperscript{3}Department of Physics, Budapest University of Technology and Economics
Several methods for community detection, but few models produce communities from microscopics.

Emergence of communities (mesoscopic structures) from microscopic mechanisms is a key question in sociology.

THIS TALK: A weighted model of (equilibrium) social networks that produces communities from microscopics.

Weights are generated dynamically and they shape the developing topology (weights $\leftrightarrow$ topology interplay).
Microscopic rules $\rightarrow$ Mesoscopic structure

$\delta = 0$

$\delta > 0$
Microscopic rules in the model

Local attachment (LA)

(1) Weighted local search / reinforcement

\[ P(i \rightarrow j) = \frac{w_{ij}}{s_i} \]
\[ P(j \rightarrow k) = \frac{w_{jk}}{(s_j - w_{ij})} \]
\[ w_{ij} \rightarrow w_{ij} + \delta \]
\[ w_{jk} \rightarrow w_{jk} + \delta \]

(2a) If (i,j,k) does not exist => Triangle formation

\[ P(i, j, k) = p_\Delta \]
\[ w_{ik} = w_0 = 1 \]

(2b) If (i,j,k) exists => Triangle reinforcement

\[ w_{ik} \rightarrow w_{ik} + \delta \]
Microscopic rules in the model

**Local attachment (LA)**

\[ k_i = 0 \implies P(i, j) = 1; w_{ij} = w_o = 1 \]

\[ k_i > 0 \implies P(i, j) = p_r; w_{ij} = w_o \]

**Global (random) attachment (GA)**

\[ k_i = 0 \implies P(i, j) = 1; w_{ij} = w_o = 1 \]

\[ k_i > 0 \implies P(i, j) = p_r; w_{ij} = w_o \]

**Node deletion (ND)**

\[ k_i > 0 \implies P(k_i = 0) = p_d \]
Microscopic rules in the model

- **Local attachment (LA)**
  
  \[ k_i > 0 \implies P(k_i = 0) = p_d \]
  
  Parameters:
  \[ w_0 = 1 \]
  \[ \delta \in [0, 1] \]

- **Global (random) attachment (GA)**
  
  \[ k_i = 0 \implies P(i, j) = 1; w_{ij} = w_o = 1 \]
  
  \[ k_i > 0 \implies P(i, j) = p_r; w_{ij} = w_o \]

- **Node deletion (ND)**
  
  \[ k_i > 0 \implies P(k_i = 0) = p_d \]

Parameters:
\[ p_\Delta, p_r, p_d \]
Initial weight
\[ w_0 = 1 \]
Increase in weight
\[ \delta \in [0, 1] \]

LA: # of links +
GA: # of links +
ND: # of links -
Microscopic rules in the model

- **Summary of model**
  - Weighted local search for new acquaintances
  - Reinforcement of existing (popular) links
  - Unweighted global search for new acquaintances

- **Parameters**
  - $\delta$: Free weight reinforcement parameter
  - $p_d = 10^{-3}$: Sets the time scale of the model $\langle \tau_N \rangle = p_d^{-1}$
  - $p_r = 5 \times 10^{-4}$: Global connections; Not sensitive
  - $p_\Delta$: Adjusted w.r.t. $\delta$ to keep $\langle k \rangle$ constant
Microscopic mechanisms in sociology

Network sociology*
- Cyclic closure
- Exponential decay
- Focal closure
- Independent of distance
- “Sample window”

Model
- Local attachment (LA)
- Global attachment (GA)
- Node deletion (ND)

Communities by inspection

- Average number of links constant $\langle L \rangle = N\langle k \rangle / 2$
  
  => All changes in structure due to reorganisation of links

- Increasing $\delta$ traps walks in communities, further enhancing trapping effect

  => Clear communities

- Triangles accumulate weight and act as nuclei for communities
Communities by k-clique method

- Use k-clique algorithm / definition for communities*
- Focus on 4-cliques (smallest non-trivial cliques)
  - Relative largest community size $R_{k=4} \in [0, 1]$
  - Average community size (excl. largest) $\langle n \rangle$
- Observe clique percolation through the system for small $\delta$
- Increasing $\delta$ leads to condensation of communities

Is community size distribution stable?

If most local random walks remain in the initial community (large $\delta$ regime), a simple argument shows that community size distribution is stationary

\[
\frac{dN_k}{dt} = -p_{d}N_k + p_{d}N \frac{N_k}{N} = 0
\]
Is community size distribution stable?

If most local random walks remain in the initial community (large $\delta$ regime), a simple argument shows that community size distribution is stationary.

\[
\frac{dN_k}{dt} = -p_d N_k + p_d N \frac{N_k}{N} = 0
\]

Rate of deleting nodes within the community
Is community size distribution stable?

If most local random walks remain in the initial community (large \( \delta \) regime), a simple argument shows that community size distribution is stationary.

\[
\frac{dN_k}{dt} = -p_d N_k + p_d N \frac{N_k}{N} = 0
\]

Rate of deleting nodes within the community

Rate at which new nodes will join the community during subsequent LA steps
Is community size distribution stable?

If most local random walks remain in the initial community (large $\delta$ regime), a simple argument shows that community size distribution is stationary.

\[
\frac{dN_k}{dt} = -p_d N_k + p_d N \frac{N_k}{N} = 0
\]

Community formation happens in transient state.

A triangle accumulating weight acts as a nucleus for the emerging community.
Weight-topology correlation

- **Weak ties hypothesis (WTH)**: The stronger the tie between nodes i and j, the greater the overlap of their friendship circles


- WTH implies weight-topology correlations: Ties within communities are strong, ties between communities are weak

- Explore weight-topology correlation with link percolation

- Control parameter $f \in [0, 1]$

- Order parameter $R_{LCC} \in [0, 1]$

* M. Granovetter, “The Strength of Weak Ties”, The American Journal of Sociology 78, 1360 (1973)
Weight-topology correlation

- Small $\delta < 0.1$
  - Network disintegrates at the same point for weak/strong link removal
  - Incompatible with WTH

- Large $\delta > 0.1$
  - Network disintegrates at different points
  - WTH compatible community structure

Weak go first  Strong go first

![Graph showing the weight-topology correlation with different thresholds](image)

alizations of $N = 5 \times 10^4$ networks. Values of $\delta$ are 0 (□), $1 \times 10^{-3}$ (∗), $1 \times 10^{-2}$ (▷), 0.1 (△), 0.5 (▽), and 1 (○).
As a model of social networks

(a) Skewed degree distribution
(b) High clustering
(c) Assortative
(d) Small world
(e) WTH compliant
Conclusion

- Model couples interaction strengths and network structure
- Communities emerge / nucleate from a structural fluctuation but only if link weight reinforcement is strong enough
- Focal closure & cyclic closure are not sufficient by themselves
- Model not only complies with the Weak Tie Hypothesis (weight-topology correlation), but suggests a plausible mechanism for it
- Suggested mechanism may be applicable to other complex networks in modelling community formation

REFERENCE: J. M. Kumpula et al., arXiv:0708.0925