A generative model for feedback networks

D.R. White¹  N. Kejžar²  C. Tsallis³  D. Farmer⁴  S. White¹

¹University of California Irvine, USA
²University of Ljubljana, Slovenia
³Centro Brasileiro de Pesquisas Físicas, Brazil
⁴Santa Fe Institute, USA

Applied Statistics, Ribno, 2005
Outline

1. Motivation
   - An example

2. Model

3. Results
   - Network properties
   - Simulations
How to model a growing network which forms cycles (establishes closer connections by adding links)?
How to model a growing network which forms cycles (establishes closer connections by adding links)?

Examples of such networks:

- **kinship network** (where to find a suitable, not blood-related, partner)
- **trading network** (search for distant trading partners to avoid the costs of paying too dearly in exchanges with close partners)
- **business network** (seeking for not too similar business partners)
How to model a growing network which forms cycles (establishes closer connections by adding links)?

Examples of such networks:

- **kinship network** *(where to find a suitable, not blood-related, partner)*

- **trading network** *(search for distant trading partners to avoid the costs of paying too dearly in exchanges with close partners)*

- **business network** *(seeking for not too similar business partners)*
How to model a growing network which forms cycles (establishes closer connections by adding links)?

Examples of such networks:

- **kinship network** *(where to find a suitable, not blood-related, partner)*
- **trading network** *(search for distant trading partners to avoid the costs of paying too dearly in exchanges with close partners)*
- **business network** *(seeking for not too similar business partners)*
Outline

1. Motivation
   - An example

2. Model

3. Results
   - Network properties
   - Simulations
An example
Creating a strategic alliance in business 3 links away.

A company which wants to make a strategic alliance.
An example
Creating a strategic alliance in business 3 links away.

Possible paths on the way. First two from the top do not lead to a successful alliance. The company chooses the link to company 3.
An example
Creating a strategic alliance in business 3 links away.

Step 1: 1 → 3

Company 3 can choose between two possible paths. The top one does not lead to a successful alliance. It chooses the link to company 6.
An example
Creating a strategic alliance in business 3 links away.

Step 2: $3 \rightarrow 6$

From company 6 there is only one way to choose the next company (company 7).
An example
Creating a strategic alliance in business 3 links away.

Step 3: $6 \rightarrow 7$

The path with 3 consecutive links was found. Alliance is created from company 1 to company 7.
Previous work

- lots of work on **generative models for graphs** (preferential attachment model of Albert and Barabási (1999), copying model of Kumar et al. (2000)); do not create cyclic networks
- **social networks model** of Newman (2003); not an evolving network model
- **autocatalytic network model** (Kauffman et al., 1986) which focused on topological graph closure properties and simulation of chemical kinetics
Previous work

- lots of work on **generative models for graphs** (preferential attachment model of Albert and Barabási (1999), copying model of Kumar et al. (2000)); do not create cyclic networks
- **social networks model** of Newman (2003); not an evolving network model
- **autocatalytic network model** (Kauffman et al., 1986) which focused on topological graph closure properties and simulation of chemical kinetics
Previous work

- lots of work on **generative models for graphs** (preferential attachment model of Albert and Barabási (1999), copying model of Kumar et al. (2000)); do not create cyclic networks
- **social networks model** of Newman (2003); not an evolving network model
- **autocatalytic network model** (Kauffman et al., 1986) which focused on topological graph closure properties and simulation of chemical kinetics
Growth of a model (1)
with 3 parameters: $\alpha, \beta, \gamma$

At each time step
- select a starting node $i$ according to probability
  \[
P_\alpha(i) = \frac{[\text{deg}(i)]^\alpha}{\sum_{m=1}^{N}[\text{deg}(m)]^\alpha}
\]
- assign of search distance $d$ according to probability
  \[
P_\beta(d) = \frac{d^{-\beta}}{\sum_{m=1}^{\infty}m^{-\beta}}
\]
- generate a search path (selection of the following nodes (l/s) on the path)
  \[
P_\gamma(l) = \frac{1 + u(l)^\gamma}{\sum_{m=1}^{M}[1 + u(m)^\gamma]}
\]

$u(x) \equiv$ unused degree of $x$
Growth of a model (1) with 3 parameters: $\alpha, \beta, \gamma$

At each time step

- **select a starting node** $i$ according to probability

  $$P_{\alpha}(i) = \frac{[\text{deg}(i)]^{\alpha}}{\sum_{m=1}^{N}[\text{deg}(m)]^{\alpha}}$$

- **assign of search distance** $d$ according to probability

  $$P_{\beta}(d) = \frac{d^{-\beta}}{\sum_{m=1}^{\infty}m^{-\beta}}$$

- **generate a search path** (selection of the following nodes ($l$s) on the path)

  $$P_{\gamma}(l) = \frac{\left[1 + u(l)^{\gamma}\right]}{\sum_{m=1}^{M}\left[1 + u(m)^{\gamma}\right]}$$

$u(x) \equiv$ unused degree of $x$
Growth of a model (1) with 3 parameters: $\alpha$, $\beta$, $\gamma$

At each time step

- select a starting node $i$ according to probability
  \[
  P_\alpha(i) = \frac{[\text{deg}(i)]^\alpha}{\sum_{m=1}^{N}[\text{deg}(m)]^\alpha}
  \]

- assign of search distance $d$ according to probability
  \[
  P_\beta(d) = \frac{d^{-\beta}}{\sum_{m=1}^{\infty}m^{-\beta}}
  \]

- generate a search path (selection of the following nodes (l's) on the path)
  \[
  P_\gamma(l) = \frac{\left[1 + u(l)^\gamma\right]}{\sum_{m=1}^{M}\left[1 + u(m)^\gamma\right]}
  \]

$u(x) \equiv$ unused degree of $x$
Growth of a model (2)

If the search path

- can be traversed for $d$ nodes, a starting node and target node are linked (a cycle is formed)
- otherwise a newly created node is linked to a starting node
Growth of a model (2)

If the search path

- can be traversed for \( d \) nodes, a starting node and target node are linked (a cycle is formed)
- otherwise a newly created node is linked to a starting node
If the search path

- can be traversed for $d$ nodes, a starting node and target node are linked (a cycle is formed)
- otherwise a newly created node is linked to a starting node

Initial condition (asymptotically not important): 1 node.
Outline

1. Motivation
   • An example

2. Model

3. Results
   • Network properties
   • Simulations
Representations of network models with 250 nodes, $\beta = 1.3$

- $\alpha = 0, \gamma = 0$
- $\alpha = 0, \gamma = 1$
- $\alpha = 1, \gamma = 0$
- $\alpha = 1, \gamma = 0$
Representations of network models with 250 nodes, $\beta = 1.3$
Role of parameters in network evolution.

\( \alpha \)\ldots the attachment parameter describes forming hubs (highly connected nodes)

\( \beta \)\ldots the distance decay parameter accounts for density of the network

\( \gamma \)\ldots the routing parameter increases search – more cycle formations, it accounts for more interconnected network
Role of parameters in network evolution.

- $\alpha$ ... the attachment parameter describes forming hubs (highly connected nodes)
- $\beta$ ... the distance decay parameter accounts for density of the network
- $\gamma$ ... the routing parameter increases search – more cycle formations, it accounts for more interconnected network
Role of parameters in network evolution.

- $\alpha$ ... the attachment parameter describes forming hubs (highly connected nodes)
- $\beta$ ... the distance decay parameter accounts for density of the network
- $\gamma$ ... the routing parameter increases search – more cycle formations, it accounts for more interconnected network
Role of parameters in network evolution.

- \( \alpha \ldots \) the attachment parameter describes forming hubs (highly connected nodes)
- \( \beta \ldots \) the distance decay parameter accounts for density of the network
- \( \gamma \ldots \) the routing parameter increases search – more cycle formations, it accounts for more interconnected network

Network evolution depends on local information, but cycle formation depends on global properties of the network:

- successful search decreases mean distance of a node to other nodes
- failed search increases the distance (with adding a new node)
Role of parameters in network evolution.

- $\alpha$... the attachment parameter describes forming hubs (highly connected nodes)
- $\beta$... the distance decay parameter accounts for density of the network
- $\gamma$... the routing parameter increases search – more cycle formations, it accounts for more interconnected network

Network evolution depends on local information, but cycle formation depends on global properties of the network:

- successful search decreases mean distance of a node to other nodes
- failed search increases the distance (with adding a new node)
Outline

1. Motivation
   - An example

2. Model

3. Results
   - Network properties
   - Simulations
Successful searches and adding nodes influence the frequency of one another → **long-range interactions among nodes**.

We simulated the networks to check whether the degree \((k)\) distributions can be described of the form (generalized \(q\)-exponential function)

\[
p(k) = p_0 k^\delta e_q^{-k/\kappa}
\]

where the **\(q\)-exponential** (Tsallis, 1988) function \(e_q^x\) is defined as

\[
e_q^x \equiv \left[ 1 + (1 - q)x \right]^{1/(1-q)} (e_1^x = e^x)
\]

if \(1 + (1 - q)x > 0\), and zero otherwise.
Simulations

The procedure

- simulate 10 realizations of networks with 5000 nodes
- different parameters $\alpha$, $\beta$ and $\gamma$
- fit generalized $q$-exponential function to simulated distributions using Gauss–Newton algorithm for nonlinear least-squares estimates (some tail regions had to be manually corrected)
- get the fitted parameters ($q$, $\kappa$ and $\delta$)
Simulations

Some results

Degree distributions and fittings for $\beta = 1.4, \gamma = 0$

Degree distributions and fittings for $\beta = 1.4, \gamma = 1$
Simulations
Goodness of fit tests

In order to test the $q$-exponential fits we used two nonparametric statistical tests

- **Kolmogorov-Smirnov test** (since $q$-exponential is defined on $[0, \infty)$ only, we used two sample test): null hypothesis was never rejected

- **Wilcoxon rank sum test**: null hypothesis rejected in 1/12 examples

Since data are very sparse in the tail, we excluded datapoints with probability $< 10^{-4}$. 
Model parameters and \( q \)-exponential

\[ p(k) = p_0 k^\delta e_{\eta}^{-k/\kappa} \]

\( \delta \) depends only on parameter \( \alpha \).
Model parameters and $q$-exponential

$p(k) = p_0 k^\gamma e^{-k/\kappa}$

Dependence of parameter $q$.

Parameter $q$ grows rapidly as each of the 3 model parameters increase.
Model parameters and $q$-exponential

\[ p(k) = p_0 k^\beta e^{-k/\kappa} \]

**Dependence of parameter $\kappa$.**

Parameter $\kappa$ diverges when $\beta$ and $\gamma$ grow large and $\alpha = 0$. 
A generative model for creating graphs representing feedback networks was presented. Algorithm uses only local properties of the nodes.

The simulated networks confirmed the assumption of long-range interactions in such a network (generalized $q$-exponential functions were fitted to empirical degree distributions).

The competition between creating cycles (stronger feedback) and adding new nodes (growth in size).

In the future

- Apply the present model to real networks (biotech intercorporate networks).
- Analyze more network model topological properties (e.g. mean distance of a node to other nodes).
Conclusion

- A generative model for creating graphs representing feedback networks was presented. Algorithm uses only local properties of the nodes.

- The simulated networks confirmed the assumption of long-range interactions in such a network (generalized q-exponential functions were fitted to empirical degree distributions).

- The competition between creating cycles (stronger feedback) and adding new nodes (growth in size).

In the future

- Apply the present model to real networks (biotech intercorporate networks).
- Analyze more network model topological properties (e.g. mean distance of a node to other nodes).
A generative model for creating graphs representing feedback networks was presented. Algorithm uses only local properties of the nodes.

The simulated networks confirmed the assumption of long-range interactions in such a network (generalized $q$-exponential functions were fitted to empirical degree distributions).

The competition between creating cycles (stronger feedback) and adding new nodes (growth in size).

In the future
- Apply the present model to real networks (biotech intercorporate networks).
- Analyze more network model topological properties (e.g. mean distance of a node to other nodes).
A generative model for creating graphs representing feedback networks was presented. Algorithm uses only local properties of the nodes.

The simulated networks confirmed the assumption of long-range interactions in such a network (generalized $q$-exponential functions were fitted to empirical degree distributions).

The competition between creating cycles (stronger feedback) and adding new nodes (growth in size).

In the future

- Apply the present model to real networks (biotech intercorporate networks).
- Analyze more network model topological properties (e.g. mean distance of a node to other nodes).
A generative model for creating graphs representing feedback networks was presented. Algorithm uses only local properties of the nodes.

The simulated networks confirmed the assumption of long-range interactions in such a network (generalized $q$-exponential functions were fitted to empirical degree distributions).

The competition between creating cycles (stronger feedback) and adding new nodes (growth in size).

In the future
- Apply the present model to real networks (biotech intercorporate networks).
- Analyze more network model topological properties (e.g. mean distance of a node to other nodes).