Probabilistic Inductive Classes of Graphs

N. Kejžar[1] V. Batagelj[2]

[1]Faculty of Social Sciences, University of Ljubljana

[2]Faculty of Mathematics, University of Ljubljana

ROeS Seminar,
Bern, Switzerland, 2007

Outline

A way to formally describe evolution of networks

- graph — a "skeleton" of the network
- model a process of graph evolution
- modelling with stochastic processes:
  - preferential attachment mechanism (Barabási, 1999)
  - copying mechanism (Kumar et al, 2000)
- inductive definition of graphs with probability transitions (transformations by rules) are viewed as implicit time steps

A way to formally describe evolution of networks

- graph — a "skeleton" of the network
- model a process of graph evolution
- modelling with stochastic processes:
  - preferential attachment mechanism (Barabási, 1999)
  - copying mechanism (Kumar et al, 2000)
- inductive definition of graphs with probability transitions (transformations by rules) are viewed as implicit time steps

A way to formally describe evolution of networks

- graph — a "skeleton" of the network
- model a process of graph evolution
- modelling with stochastic processes:
  - preferential attachment mechanism (Barabási, 1999)
  - copying mechanism (Kumar et al, 2000)
- inductive definition of graphs with probability transitions (transformations by rules) are viewed as implicit time steps

A way to formally describe evolution of networks

- graph — a "skeleton" of the network
- model a process of graph evolution
- modelling with stochastic processes:
  - preferential attachment mechanism (Barabási, 1999)
  - copying mechanism (Kumar et al, 2000)
- inductive definition of graphs with probability transitions (transformations by rules) are viewed as implicit time steps

Inductive class of graphs

Inductive class of graphs

ICG is defined as $\mathcal{I} = (\mathcal{B}; \mathcal{R})$ (Curry, 1963):

1. class $\mathcal{B}$ of initial graphs, the basis of ICG,
2. class $\mathcal{R}$ of generating rules, which consists of a distinguished **left element** (part of a graph) to which the rule is applied (the left element is transformed into the **right element** of the rule).

Inductive class of graphs

Inductive class of graphs

ICG is defined as $\mathcal{I} = (\mathcal{B}; \mathcal{R})$ (Curry, 1963):

1. class $\mathcal{B}$ of initial graphs, the basis of ICG,
2. class $\mathcal{R}$ of generating rules, which consists of a distinguished **left element** (part of a graph) to which the rule is applied (the left element is transformed into the **right element** of the rule).

The inductive class consists exactly of the graphs that can be obtained from the basis in finite number of steps using the generating rules.

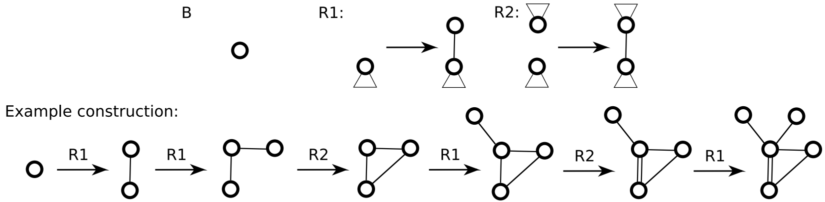
Inductive class of graphs

Nice properties of generating rules:

- locality (the part of the graph, where to apply the rule is connected)
- expansion (the rule increases a selected property of graphs; i.e. number of vertices)

Example

$$\mathcal{I} = (\mathcal{B}; \mathcal{R}), \mathcal{B} = \{B\}, \mathcal{R} = \{R1, R2\}$$



Probabilistic ICG

Probabilistic inductive class of graphs

PICG is defined as $\mathcal{P} = (\mathcal{B}; \mathcal{R}; f_B; f_R; \mathcal{F}_L)$:

1. class $\mathcal{B}$ of initial graphs, the basis,
2. class $\mathcal{R}$ of generating rules
3. probability distribution f_B specifying how the initial graph is chosen from class $\mathcal{B}$,
4. probability distributions f_R specifying how the rules from class $\mathcal{R}$ are applied,
5. a set of $|\mathcal{R}|$ probability distributions — $\mathcal{F}_L$ specifying how the left elements for every rule in class $\mathcal{R}$ are chosen.

Probabilistic ICG

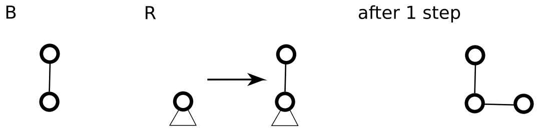
Probabilistic inductive class of graphs

PICG is defined as $\mathcal{P} = (\mathcal{B}; \mathcal{R}; f_B; f_R; \mathcal{F}_L)$:

1. class $\mathcal{B}$ of initial graphs, the basis,
2. class $\mathcal{R}$ of generating rules
3. probability distribution f_B specifying how the initial graph is chosen from class $\mathcal{B}$,
4. probability distributions f_R specifying how the rules from class $\mathcal{R}$ are applied,
5. a set of $|\mathcal{R}|$ probability distributions — $\mathcal{F}_L$ specifying how the left elements for every rule in class $\mathcal{R}$ are chosen.

We limit on simple PICG definitions in which the left part of the selected rule can always be found.

Preferential attachment model as PICG



$\mathcal{P} = (\mathcal{B}; \mathcal{R}; f_B; f_R; \mathcal{F}_L):, \mathcal{B} = \{B\}, \mathcal{R} = \{R\}$

- $f_B = 1$
- $f_R = 1$
- $f_u = \frac{deg(u)}{\sum_v deg(v)}, \qquad \forall u \in V, f_u \in \mathcal{F}_L$

Expected number of vertices — order

Expected change in the number of vertices has to be positive (otherwise graph dies out).
Probability for number of vertices (recursive expression):

$$p_t(N = n) = \sum_{R_i \in \mathcal{R}} r_i p_{t-1}(N = n - \Delta n_i)$$

- N — random variable for number of vertices
- t — time step
- $r_i \equiv f_R(R_i)$ — probability of selecting rule R_i
- Δn_i the number of vertices, that rule R_i adds to the graph

Initial values for graphs from the basis:

$$p_0(N = n) = \sum_{B_i; |B_i| = k} f_B(B_i)$$

Expected change in number of vertices

Expected order in time t:

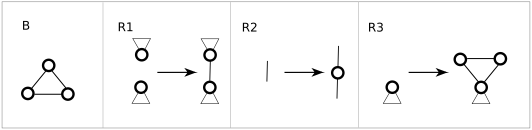
$$E_t[N] = \sum_{i=0}^{\infty} i p_t(i)$$

It can be shown, that in the limit of large t:

$$\lim_{t\to\infty} \frac{E_t[N]}{t} = \sum_{R_i \in \mathcal{R}} r_i \Delta n_i$$

Similarly, this holds for number of edges.

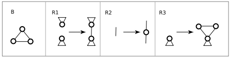
Application — 2-edge-connected graphs



$$\mathcal{P}_{2E} = (B; \{R1, R2, R3\}; 1; f_R, \mathcal{F}_L)$$

- $f_R = \begin{pmatrix} R1 & R2 & R3 \\ q & r & s \end{pmatrix}$
- $\mathcal{F}_L = \{f_{R1}, f_{R2}, f_{R3}\}$
  - $f_{R1}(u, v) = \frac{1}{n(n-1)}$
  - $f_{R2}(e) = \frac{1}{m}$
  - $f_{R3}(u) = \frac{1}{n}$

... 2-edge-connected graphs



From the recursive relation:

$$p_t(n) = qp_{t-1}(n) + rp_{t-1}(n-1) + sp_{t-1}(n-2)$$

we get

$$p_t(n) = \sum_{j=\lfloor \frac{n}{2} \rfloor}^{n-2} \binom{j-1}{n-j-2} \binom{t}{t-j+1} q^{t-j+1} r^{2j-n+1} s^{n-j-2}$$

Therefore expected change in vertices:

$$\frac{E_t[N]}{t} \approx r + 2s$$

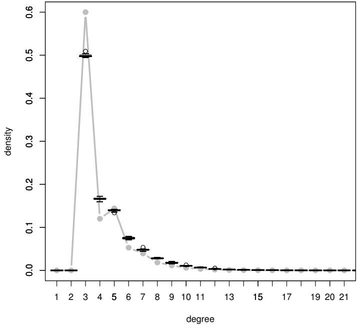
Similarly — expected change in edges:

$$\lim_{t \to \infty} \frac{E[M]}{t} = (q+r) + 3s$$

... 2-edge-connected graphs

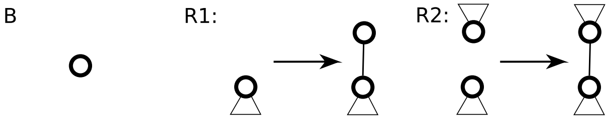
Degree distribution (p_k — probability for a vertex to have a degree k):

$$(n+r+2s)p'_k = np_k + r\left(\delta_{k2}\right) + q\left(p_{k-1} - p_k\right) + s\left(p_{k-2} + 2\delta_{k2} - p_k\right)$$



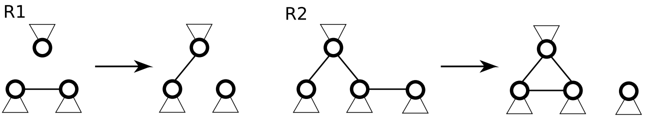
Spread of rumors

- basis — person with "information"
- $R1$ — a person with a rumor disseminates information to a person without
- $R2$ — two people with knowledge of a rumor discuss it



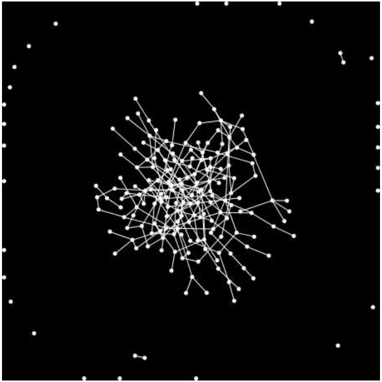
Dynamics of acquaintanceships

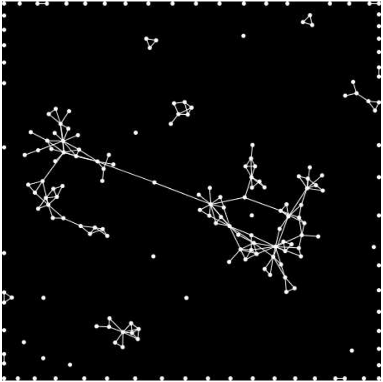
- basis — any graph with n vertices and m edges (i.e. ER random graph)
- $R1$ — breaking up of an acquaintanceship for a more "advantageous" one
- $R2$ — creation of "stronger" relationships by forming connected triads (friendship)



Dynamics of acquaintanceships

Left: basis; ER model, 200 vertices and 200 edges
Desno: graph over 1000 steps ($p_1 = 0.11$ in $p_2 = 0.89$)





Dynamic of social balance processes

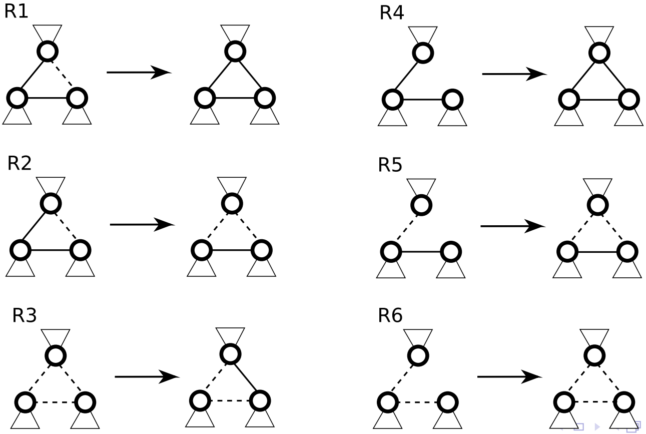
Structural balance theory (Cartwright, Harary 1956): signed edges between two vertices

Balanced triads: signs of edges among all three vertices multiply into a positive sign

People tend to gravitate towards balanced triads with other people.

Dynamic of social balance processes

- basis — simple graph with n vertices and m signed edges
- $R1, R2, R3$ — balancing rules
- $R4, R5, R6$ — process of induction (Heider 1946)



Conclusion

- PICG is a framework which enables a formal description of many network models.
- adding and deleting of properties of edges and vertices can be included
- more results about some formal models (i.e. degree distribution) can be obtained at `arXiv:math/0612778v1 [math.DS]`

Conclusion

- PICG is a framework which enables a formal description of many network models.
- adding and deleting of properties of edges and vertices can be included
- more results about some formal models (i.e. degree distribution) can be obtained at `arXiv:math/0612778v1 [math.DS]`

Conclusion

- PICG is a framework which enables a formal description of many network models.
- adding and deleting of properties of edges and vertices can be included
- more results about some formal models (i.e. degree distribution) can be obtained at `arXiv:math/0612778v1 [math.DS]`