Novel Models of Group Selection in Social Structures and Networks

David Hales

University of Bologna, Italy

www.davidhales.com

In collaboration with: Stefano Arteconi, Andrea Marcozi, Simon Patarin, Ozalp Babaoglu

Distributed computer systems are making new kinds of social systems. By engineering them in certain ways we change social realities rather than merely trying to reflect them.

Social science and distributed systems engineering are merging.

Overview

- What are self-* systems?
- Peer-to-Peer (P2P) systems
- The BitTorrent file-sharing system
- Recent group selection models
- Group selection in P2P
- Note on methods

What is Self-*

- Information systems that
  - Self-organise
  - Self-manage
  - Self-repair
  - Self-adapt
- Without explicit administrative or user intervention

What is Self-*

- New trend in information systems research because increasingly:
  - Open distributed systems
  - Without central control
  - Massive (millions of components)
  - Dynamic and noisy (at run time)
  - Standard design approaches fail

Technology areas in Self-*

- Grids, MAS
- Ad hoc networks (mob. phones, PDA's)
- Autonomic systems (top-down) self-adaptive
- Peer-to-Peer (P2P) systems

Recent new conference

- SASO: Self-Adaptive and Self-Organising Systems
- IEEE sponsored
- Merger of ESOA, SelfMan, Self-* and IWSAS workshops
- First one July 2007 @ MIT
- http://projects.csail.mit.edu/saso2007/

Peer-to-Peer Systems

What are P2P systems?

- Machines (nodes) on the internet
- Dynamically connecting to a few others
- Cooperating to achieve some task
- So-called "overlay networks"
- Majority of internet bandwidth use is P2P today
- Often associated with illegal copying

Popular applications of P2P

- BitTorrent
  - Open protocol for sharing large files
  - Peers cooperate to speedup downloads
- Skype
  - Closed protocol for voice over IP
  - Peers cooperate to route audio streams
- Joost (beta)
  - Internet based TV

What has this got to do with social simulation?

- P2P need algorithms that are:
  - Decentralised (no central control)
  - Scalable (to millions)
  - Robust (to failure, noise, and malicious)
  - Simple (lightweight code)
  - Promote cooperation (avoid free-riding)
- Isn't this what a lot of algorithms from social simulation do?

Social Simulation Contributions to P2P

- Social simulation work can contribute in two distinct ways:
  - Supply algorithms for implementations
  - Supply "user models" which capture how users interact with systems
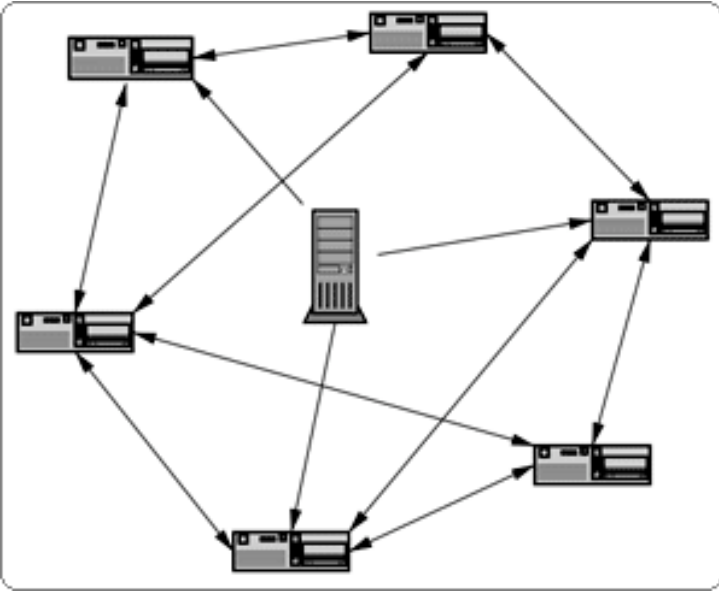- I will mainly focus on the first of these today

Overview of BitTorrent

- Most popular file-sharing P2P protocol
- Peers cooperatively pool resources
- Open protocol so anyone can write their own "peer client" software
- Based on the tit-for-tat cooperation strategy popularised by Robert Axelrod
- Creator: Bram Cohen

BitTorrent
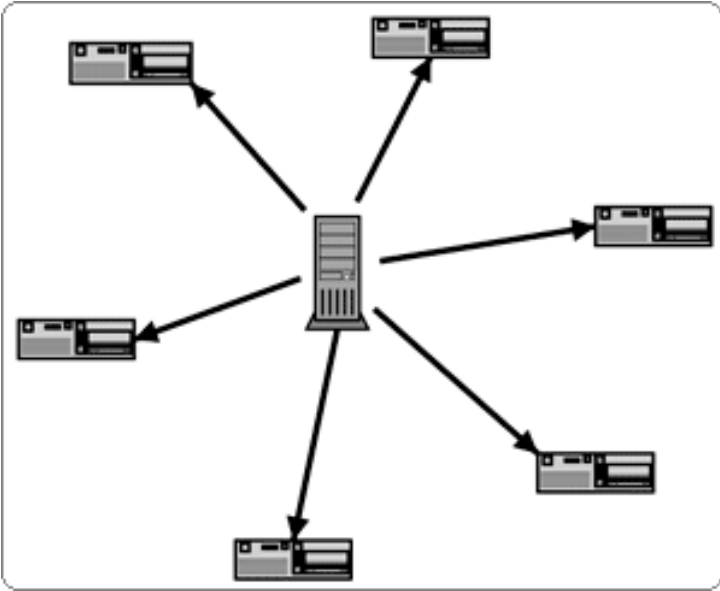
Central Server Approach   BitTorrent Approach
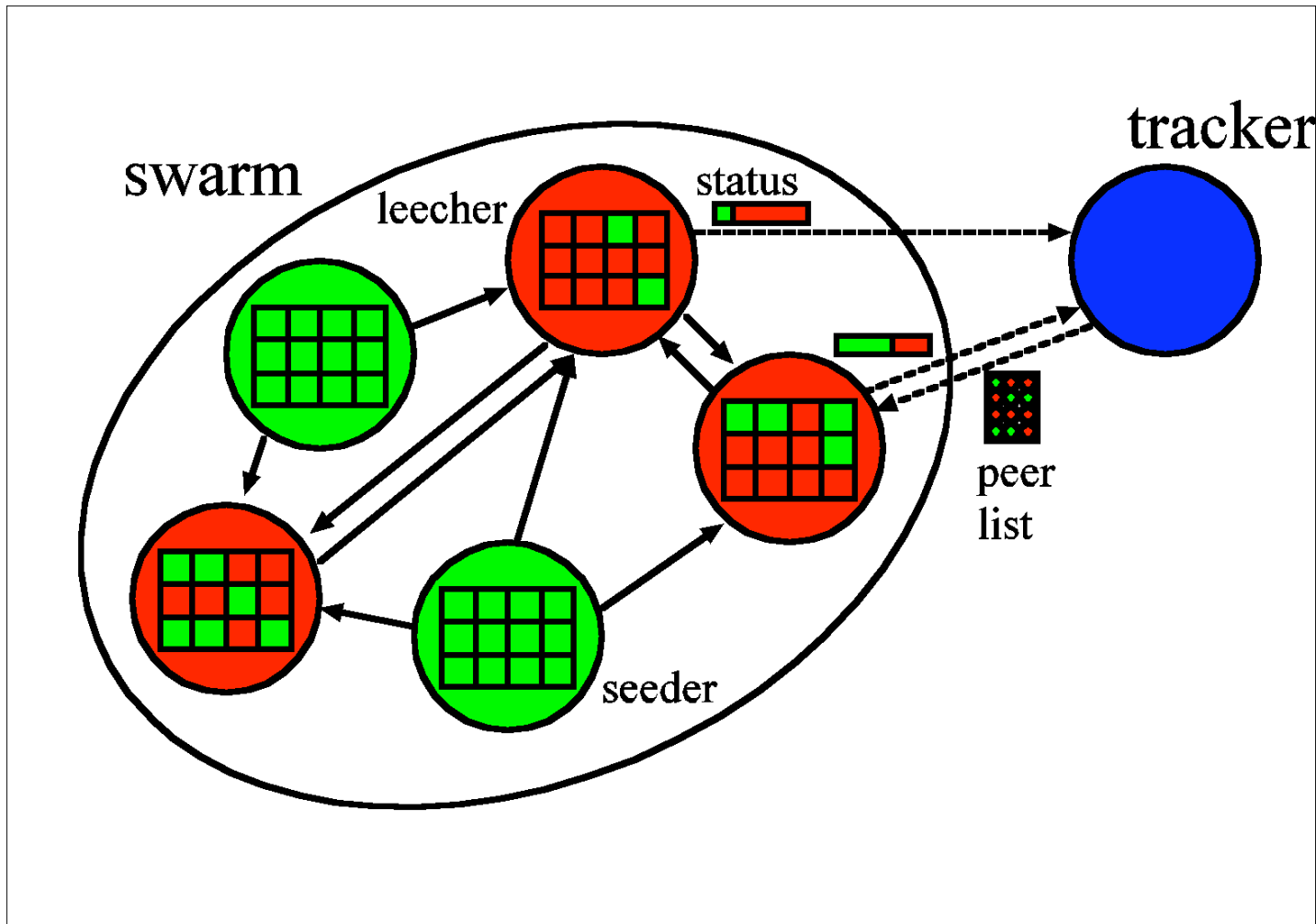
BitTorrent

- When a node wishes to share a file it:
    - splits it into many small chunks
    - creates a new "swarm" containing itself
    - publishes a pointer (.torrent file) to the swarm
- To download a file a node:
    - uses the .torrent file to join the associated swarm
    - connects to several other nodes in the swarm
    - downloads the blocks it needs
    - uploads requested blocks to others

BitTorrent

- While downloading nodes
  - Monitor performance of each link
  - Drop links when uploading is not being reciprocated
  - Keep links which are reciprocating
  - Occasionally try new random links

BitTorrent

BitTorrent

- This is a kind of tit-for-tat strategy
- Cooperation = upload to others
- Defection = only download from others
- By breaking links to selfish nodes (so called leechers) free-riding not viable
- If you don't upload you don't download

Bad guys strike back!

http://bittyrant.cs.washington.edu/

BitTyrant

A strategic BitTorrent client that improves performance

BitThief   http://dcg.ethz.ch/projects/bitthief/

A Free Riding BitTorrent Client

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Bad guys strike back!

- BitTorrent can still be cheated
- Selfish clients have been released by researchers to see if they spread
- BitTorrent is becoming a *global social cooperation experiment*
- The jury is still out on why selfish clients do not seem to have taken over
- Game theorists seem to be confused

BBC   Home   News   Sport   Radio   TV   Weather   Languages   Search

UK version   International version | About the versions   Low graphics | Accessibility help

BBC NEWS   WATCH One-Minute World News   News services
Your news when you want it

Last Updated: Friday, 31 August 2007, 10:21 GMT 11:21 UK

E-mail this to a friend   Printable version

File-sharers forced to play fair

By Colin Barras
Technology reporter

Researchers have found a way to enforce good manners on file-sharing networks by treating bandwidth as a currency.

The team has created a peer-to-peer system called Tribler in which selfless sharers earn faster upload and download speeds but leechers are penalised.

Tribler has also been made to work with the PlayStation 3

The technology is being assessed by a European broadcasting body looking at ways of piping TV across the net.

Tribler has also been used to turn Sony's PlayStation 3 into a video-sharing device.

Fair sharing

While file-sharing networks are good ways to help lots of

SEE ALSO
▶ Speed boost plan for file-sharing
12 Apr 07 | Technology
▶ UK net users leading TV downloads
19 Feb 05 | Technology
▶ Troubled times for home networks
21 Aug 07 | Technology
▶ Net firm warns on web video costs
13 Aug 07 | Technology
▶ Warnings of 'internet overload'
15 Jun 07 | Click
▶ PlayStation to record digital TV
22 Aug 07 | Technology
▶ BBC online video service launches
27 Jul 07 | Technology

RELATED INTERNET LINKS
▶ Tribler project
▶ PlayStation
The BBC is not responsible for the content of external internet sites

TOP TECHNOLOGY STORIES

Done

Files

Video Files   Name   search files, YouTube

035 BAYROU ET   133 MACHINES A   142 NEW VOTE   169_battlefield   280407_vanity.mov
L'UNIVERSELLE.mov   VOTER .mov   ELECTRONIQUE   bad_ot
                                      N.mov

300_theatrical.mov 35335403.mov   AMV - Ghost In The   Card.Game.2005.Xvid Crysis HDTV
                                      Shell - Vectorman   Trailer.wmv

Jeremy.mov   Kns Kali - We're   Mustache Maddness - NC_AUTOEXE_RXBSE   POTC3_720i.mov
                   Missing You.mpg   episode one.mov

Pow-wow Drumming   Resident_evil_extinct S01_E06   SFYouTube720pWW.m   SILVER_SURFER_HD.z
01.AVI

◄ 1 2 3 ►

Resident_evil_extinction.divx

info   Files

size: 21.12 MB
creation date: 25-02-2007
popularity: 129   1
fit to taste:

Description

Resident Evil Extinction trailer

People who like this also like

No similar files found yet.

Downloading (0)   down: 0 KB/s | up: 0 KB/s

- **Wealth of Content:**
  BitTorrent, YouTube, and Liveleak

- **All-in-one Solution:**
  Serverless P2P File Sharing, VLC player,
  Divx codecs, Flash plugin, Podcasts, ...

- **Intelligent Taste Recognition**

- **Go beyond BitTorrent...**

It's easy: select a version, download, install, and start watching

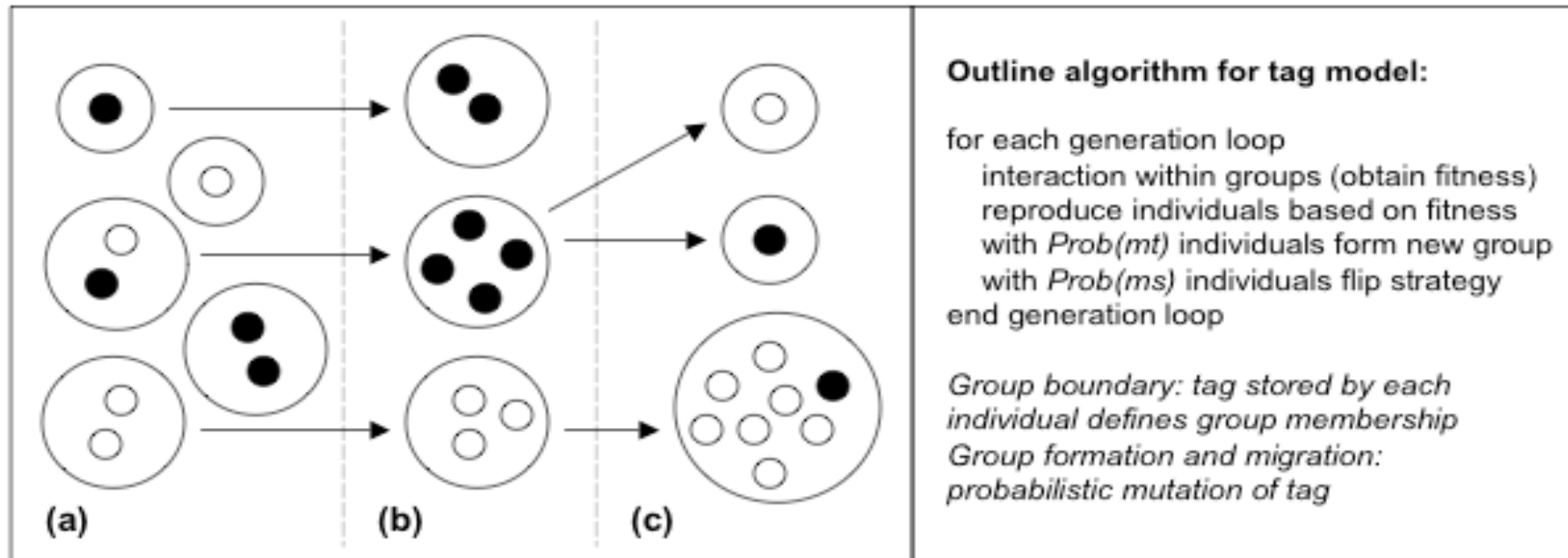| ●	●
• **Minimize your upload to others** • **Download videos 20% faster**	• **Upload as much as you download** • **Download videos at normal speed**

DOWNLOAD TRIBLER
for Windows

Done

New Group Selection Models

Group Selection Models

- Recent models of "group selection"
- Based on individual selection
- Producing dynamic social structures
- Limit free-riding
- Increasingly group-level performance
- Don't require reciprocity
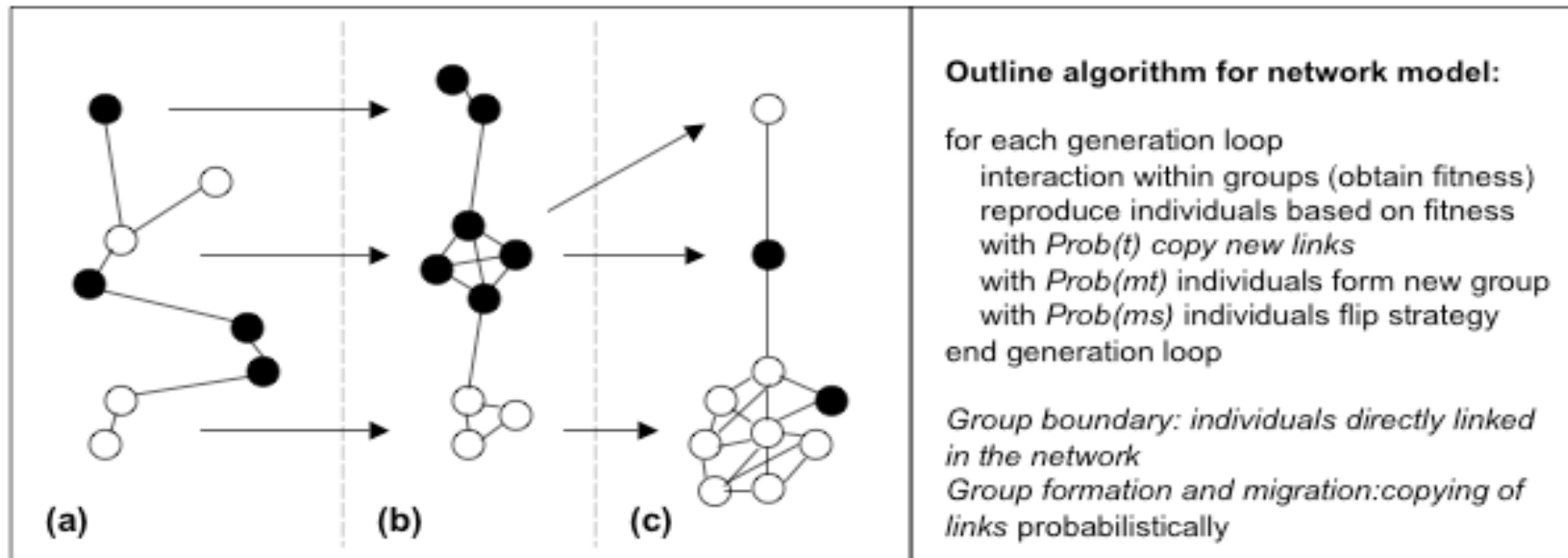- Could be very useful in P2P

Evolutionary Group Selection Models

- *Group boundary* - a mechanism which restricts interactions between agents such that the population is partitioned into groups

- *Group formation* - a process which forms groups dynamically in the population

- *Migration* - a process by which agents may move between different groups

- *Conditions* - cost / benefit ratio of individual interactions and other conditions which are sufficient for producing group-level selection

Outline algorithm for tag model:

for each generation loop
    interaction within groups (obtain fitness)
    reproduce individuals based on fitness
    with *Prob(mt)* individuals form new group
    with *Prob(ms)* individuals flip strategy
end generation loop

Group boundary: tag stored by each individual defines group membership
Group formation and migration: probabilistic mutation of tag

(a)   (b)   (c)

Schematic of the evolution of groups in the tag model. Three generations (a-c) are shown. White individuals are pro-social (altruistic), black are selfish. Individuals sharing the same tag are shown clustered and bounded by large circles. Arrows indicate group linage. When **b** is the benefit a pro-social agent can confer on another and **c** is the cost to that agent then the condition for group selection of pro-social groups is: **b > c and mt >> ms**
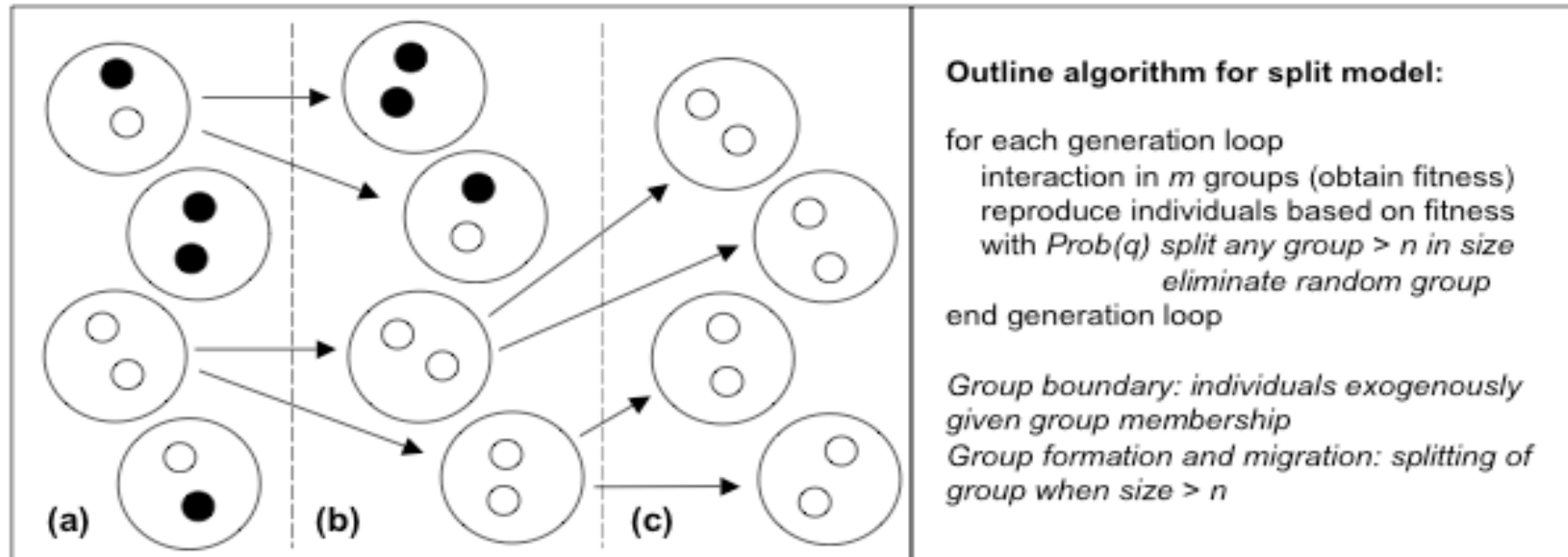
Riolo, Axelrod, Cohen, Holland, Hales, Edmonds…

Outline algorithm for network model:

for each generation loop
    interaction within groups (obtain fitness)
    reproduce individuals based on fitness
    with *Prob(t)* copy new links
    with *Prob(mt)* individuals form new group
    with *Prob(ms)* individuals flip strategy
end generation loop

Group boundary: individuals directly linked in the network
Group formation and migration:copying of links probabilistically

Schematic of the evolution of groups in the network-rewire model. Three generations (a-c) are shown. Altruism selected when:**b > c and mt >> ms.** When **t = 1**, get disconnected components, when **1 > t > 0.5**, get small-world networks

Hales, D. & Arteconi, S. (2006) Article: SLACER: A Self-Organizing Protocol for Coordination in P2P Networks. IEEE Intelligent Systems, 21(2):29-35

Santos F. C., Pacheco J. M., Lenaerts T. (2006) Cooperation prevails when individuals adjust their social ties. PLoS Comput Biol 2(10)

Outline algorithm for split model:

for each generation loop
  interaction in *m* groups (obtain fitness)
  reproduce individuals based on fitness
  with *Prob(q)* split any group > n in size
        eliminate random group
end generation loop

Group boundary: individuals exogenously given group membership
Group formation and migration: splitting of group when size > n

Schematic of the evolution of in the group-splitting model. Three generations (a-c) are shown. Altruism is selected if the population is partitioned into *m* groups of maximum size *n* and *b / c > 1 + n / m.*

Traulsen, A. & Nowak, M. A. (2006). Evolution of cooperation by multilevel selection. Proceedings of the National Academy of Sciences 130(29):10952-10955.

SLAC: Network re-wire P2P model

- Agents = nodes in a P2P overlay network
- Each node links to some neighbors (view) in overlay
- Assume:
  - Interaction between neighbors to achive some application task
  - Behavior: Application behavior (i.e. share files or leech files, cooperate or defect)
  - Utility: Evaluated at application level (i.e. number of files downloaded, performace metric)

SLAC algorithm

Each node p periodically executes the following:

q = SelectRandomPeer()
if utility$_q$ > utility$_p$

      drop all current links

      link to node q and copy its strategy and links

      mutate (with low probability) strategy and links
fi
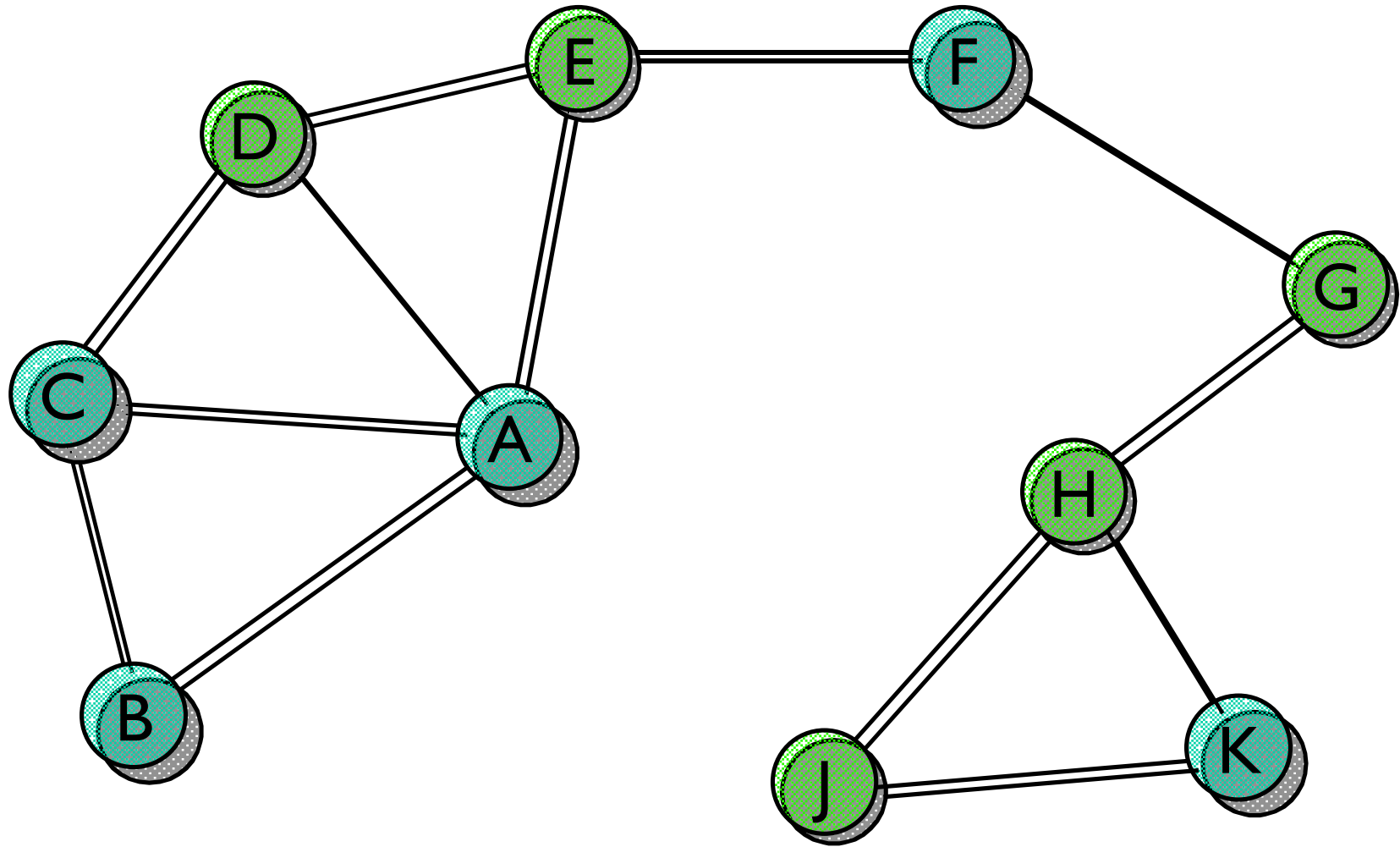
SLAC: "Copy and Rewire"

SLAC: "Copy and Rewire"

Compare utilities

SLAC: "Copy and Rewire"

"Copy" strategy

SLAC: "Copy and Rewire"

Drop current links
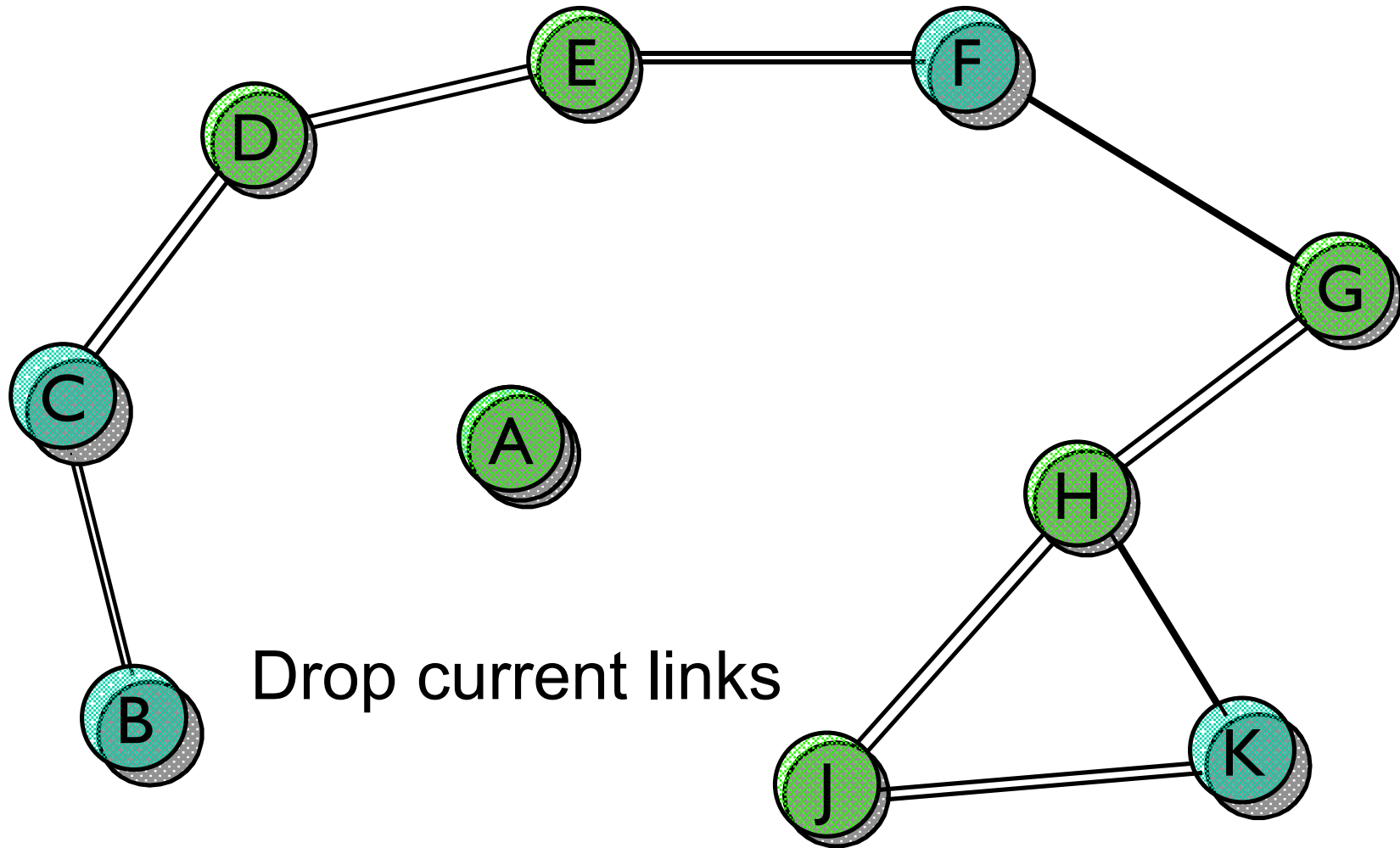
SLAC: "Copy and Rewire"

"Rewire"
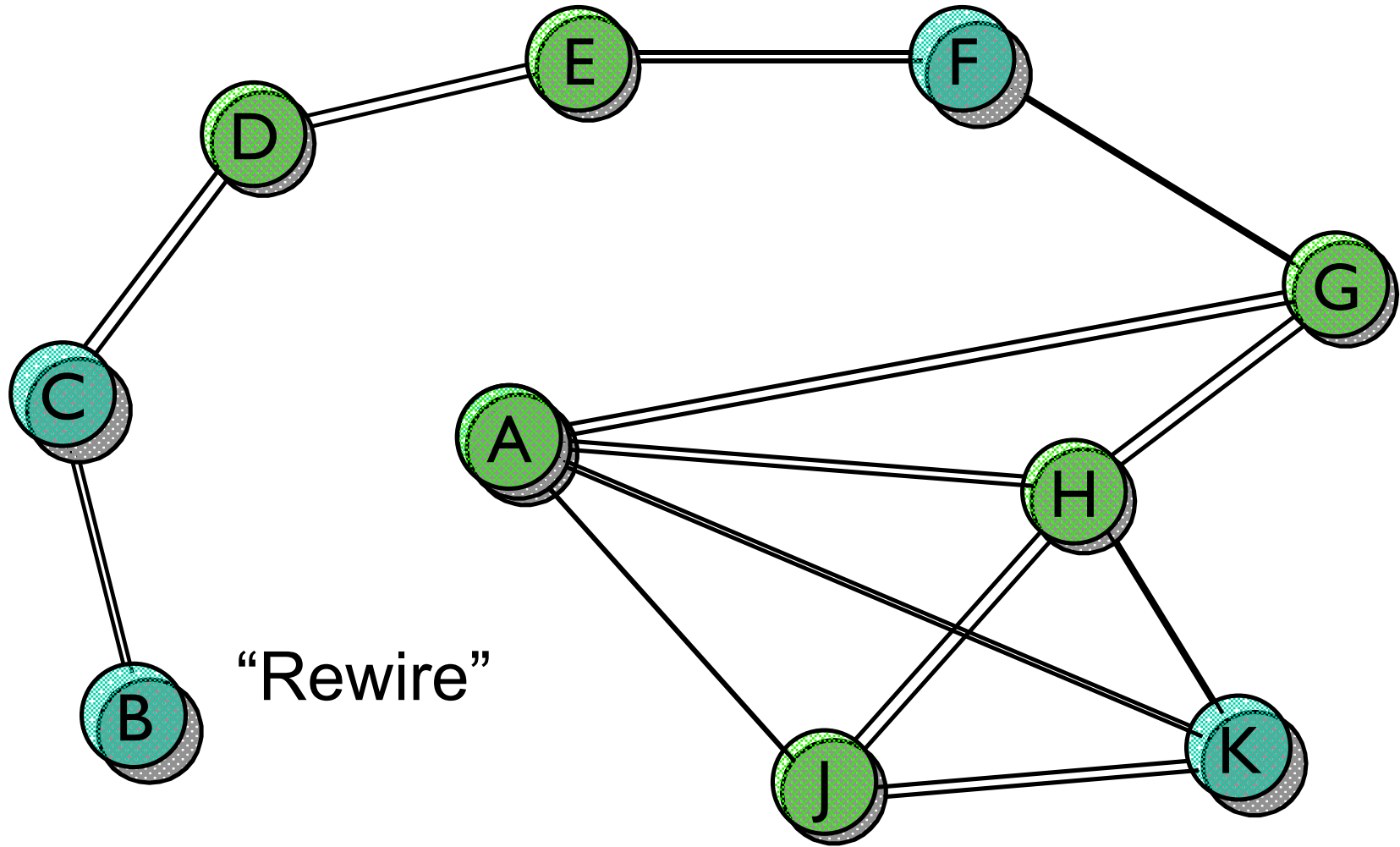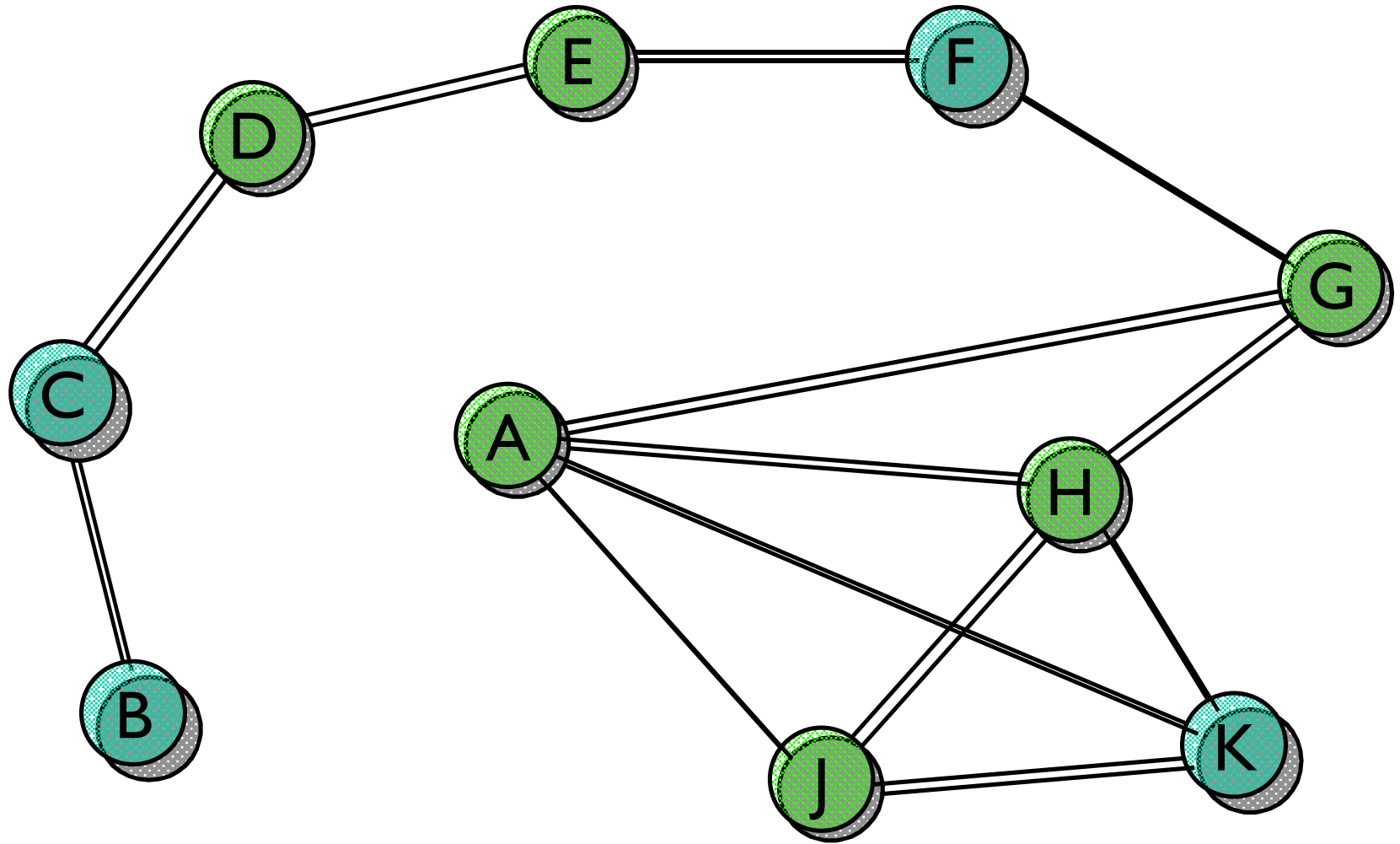
SLAC: "Mutate"

SLAC: "Mutate"

"Mutate" strategy

SLAC: "Mutate"


Drop current links
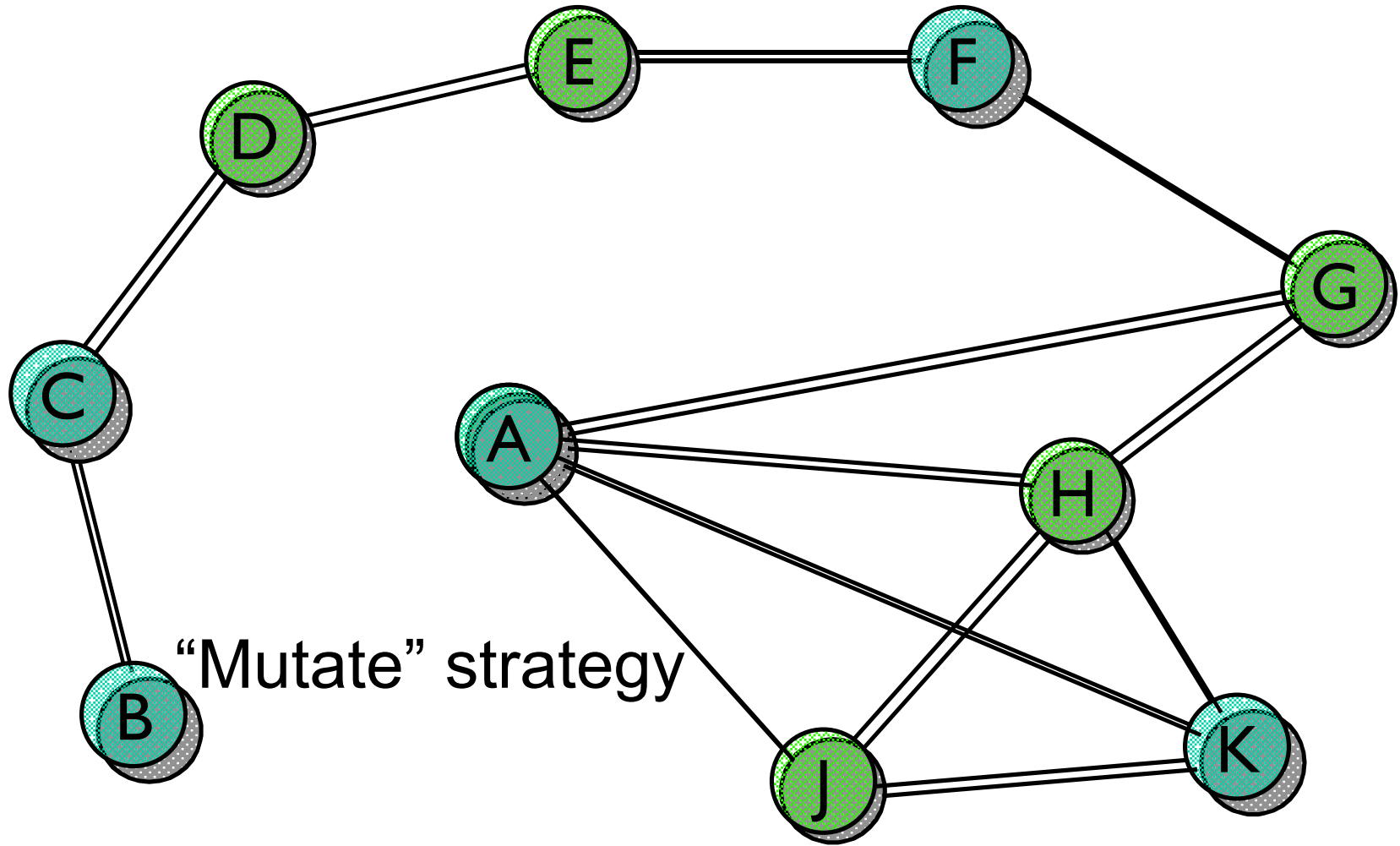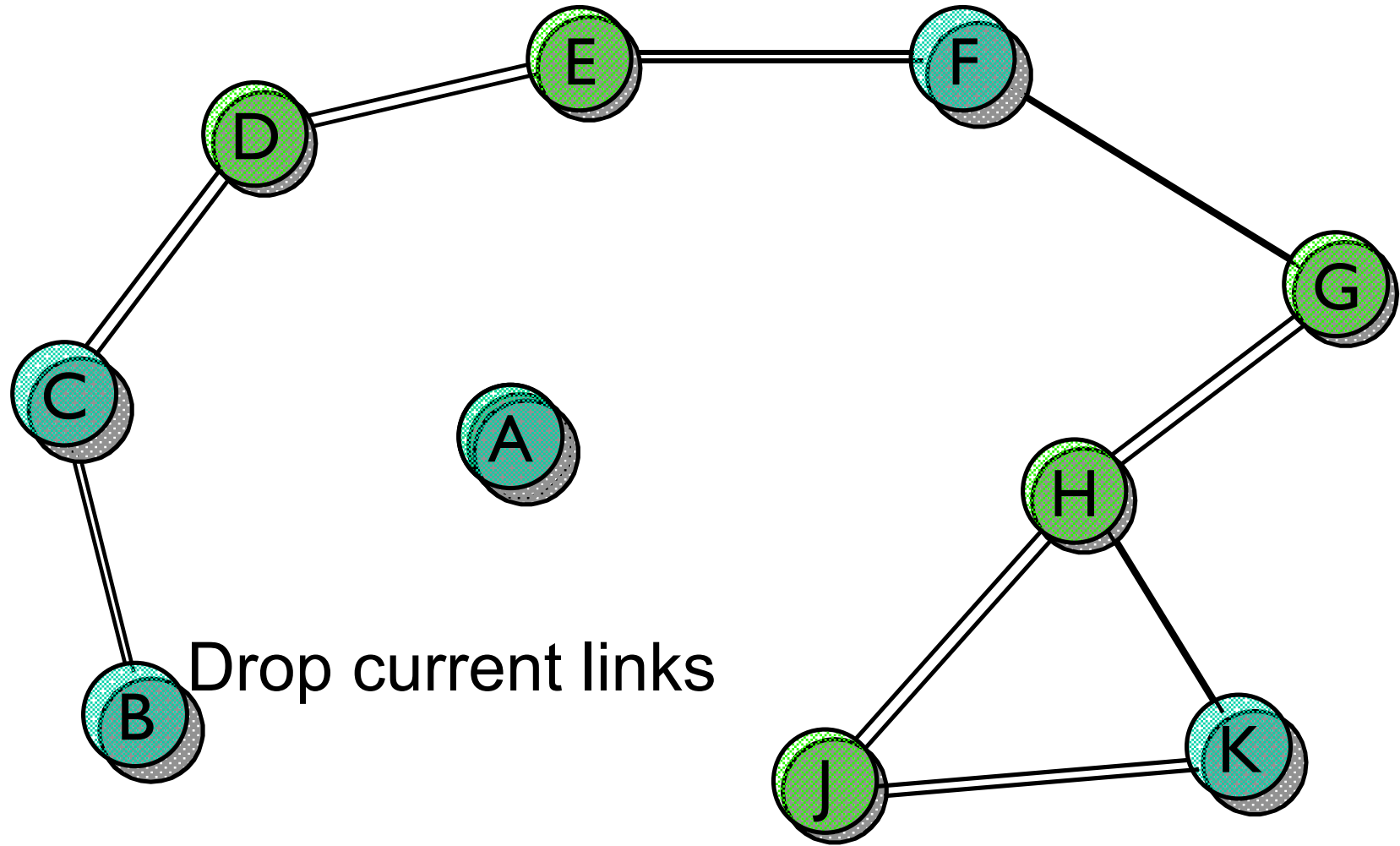
SLAC: "Mutate"

Link to random node
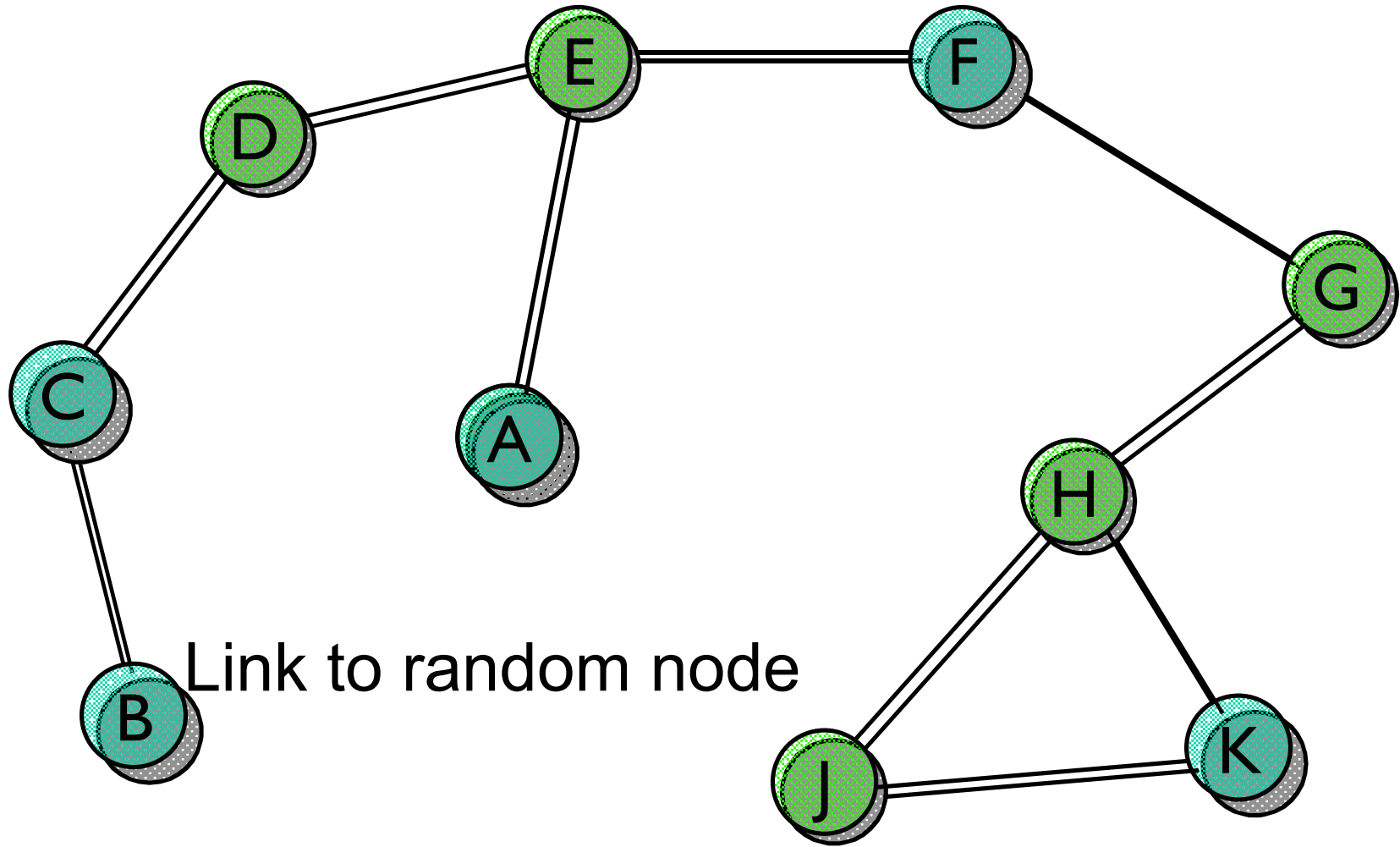
SLAC playing the PD

- We tested SLAC with Prisoner's Dilemma (PD)
  - Captures the conflict between "individual rationality" and "common good"
  - Defection (D) leads to higher *individual* utility
  - Cooperation (C) leads to higher *global* utility
  - $DC > CC > DD > CD$
- Prisoner's Dilemma in SLAC
  - Nodes play PD with neighbors chosen randomly in the interaction network
  - Only pure strategies (always C or always D)
  - Strategy mutation: flip current strategy
  - Utility: average payoff achieved

Cycle 180: Small Defect Clusters
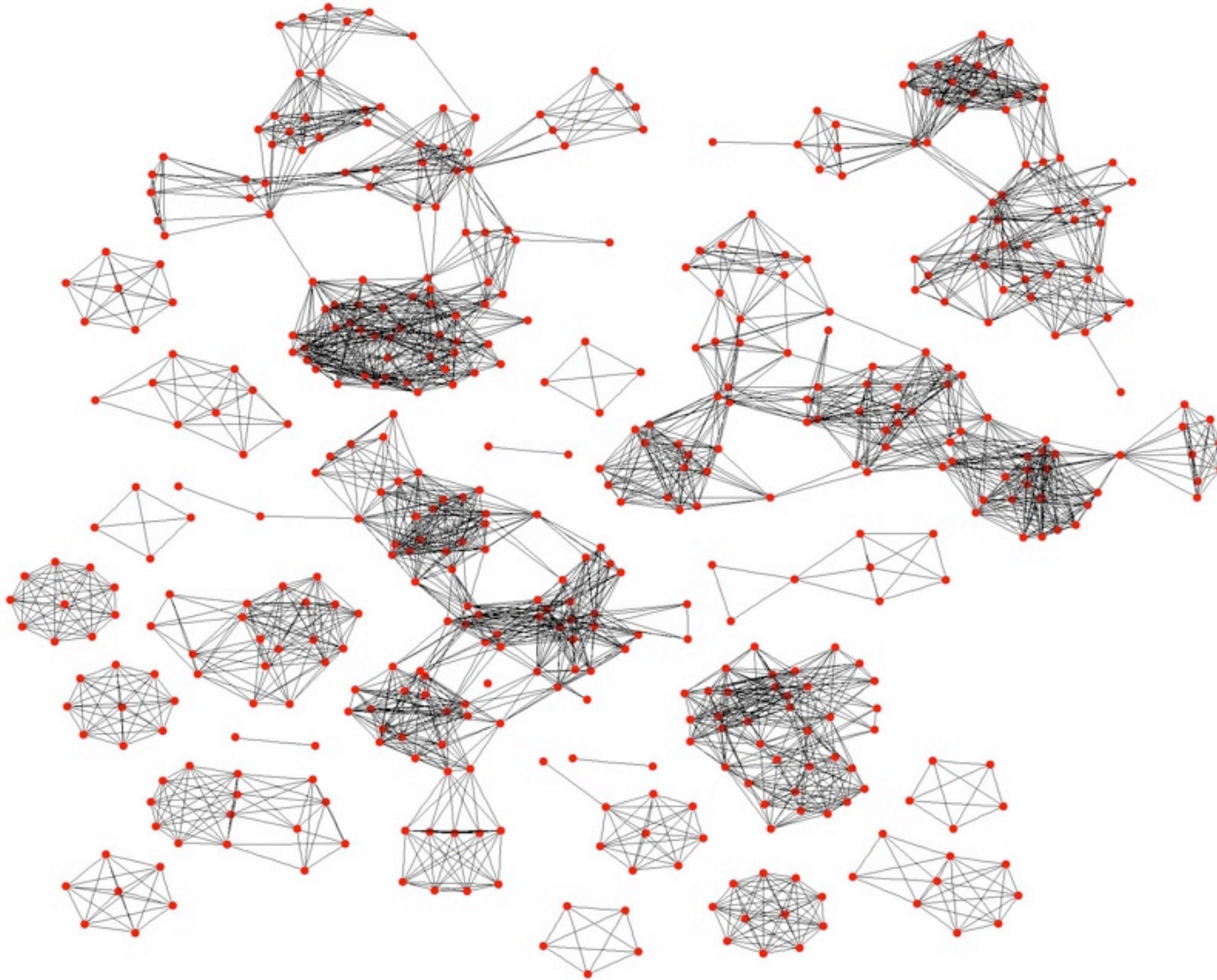
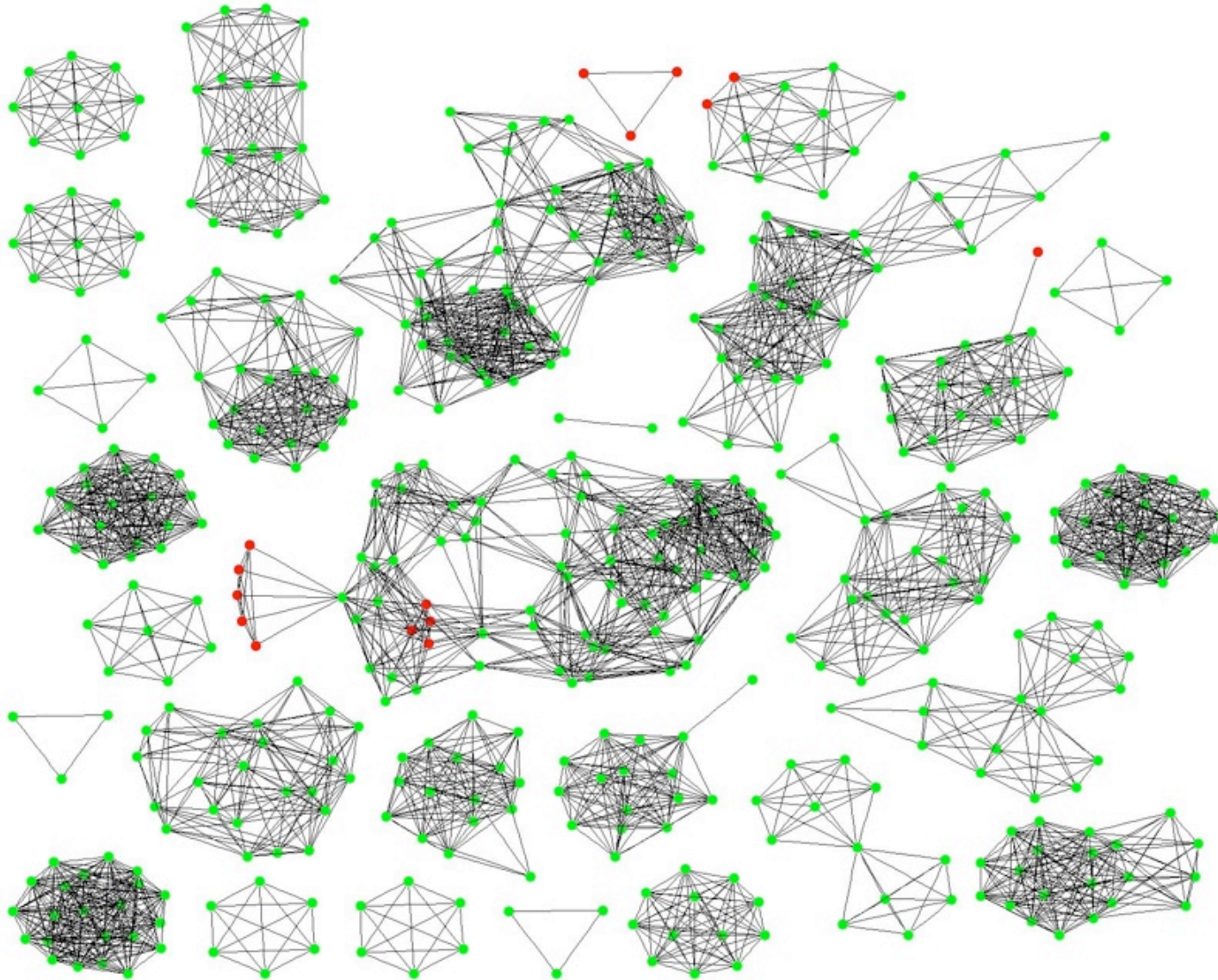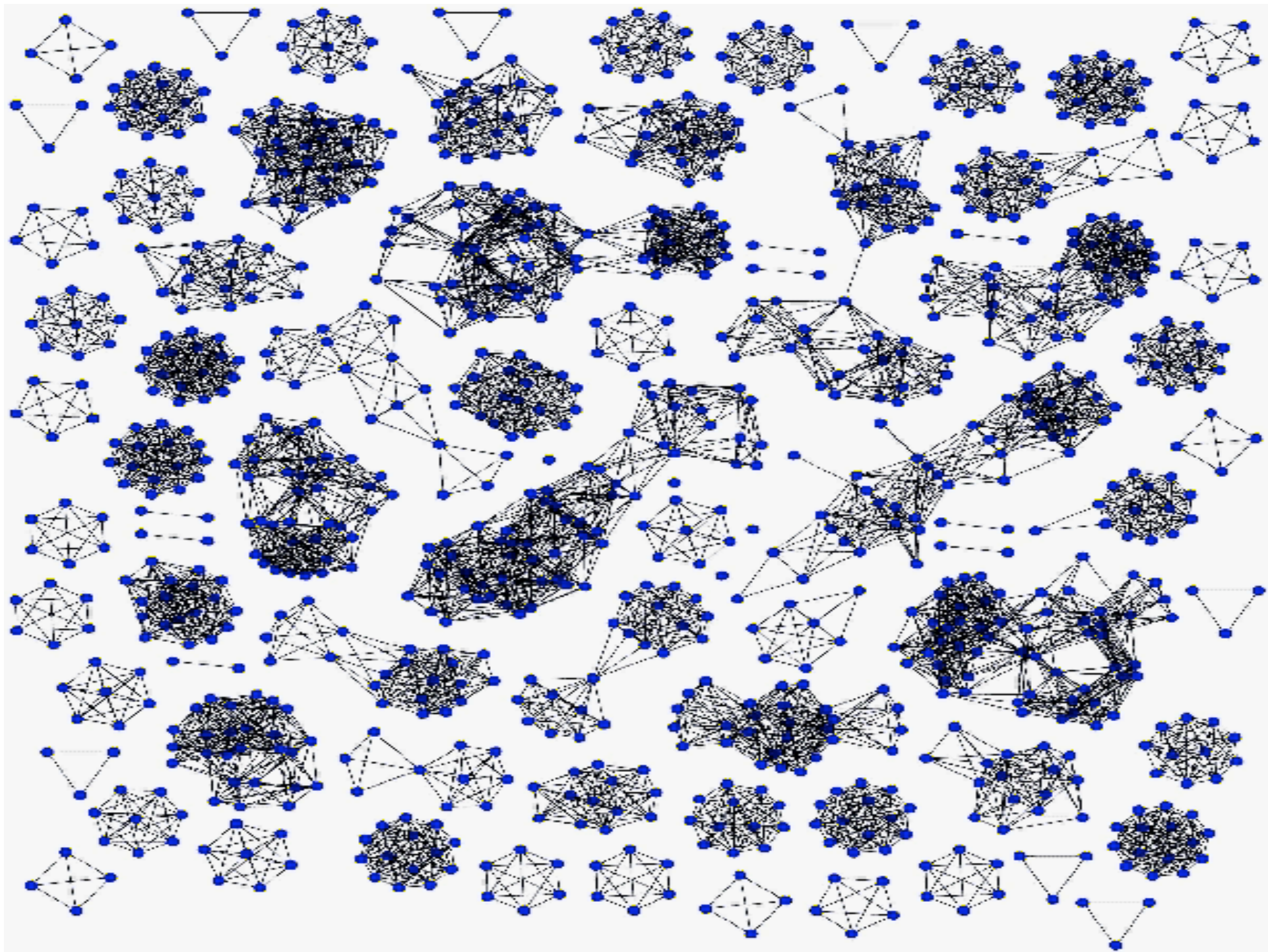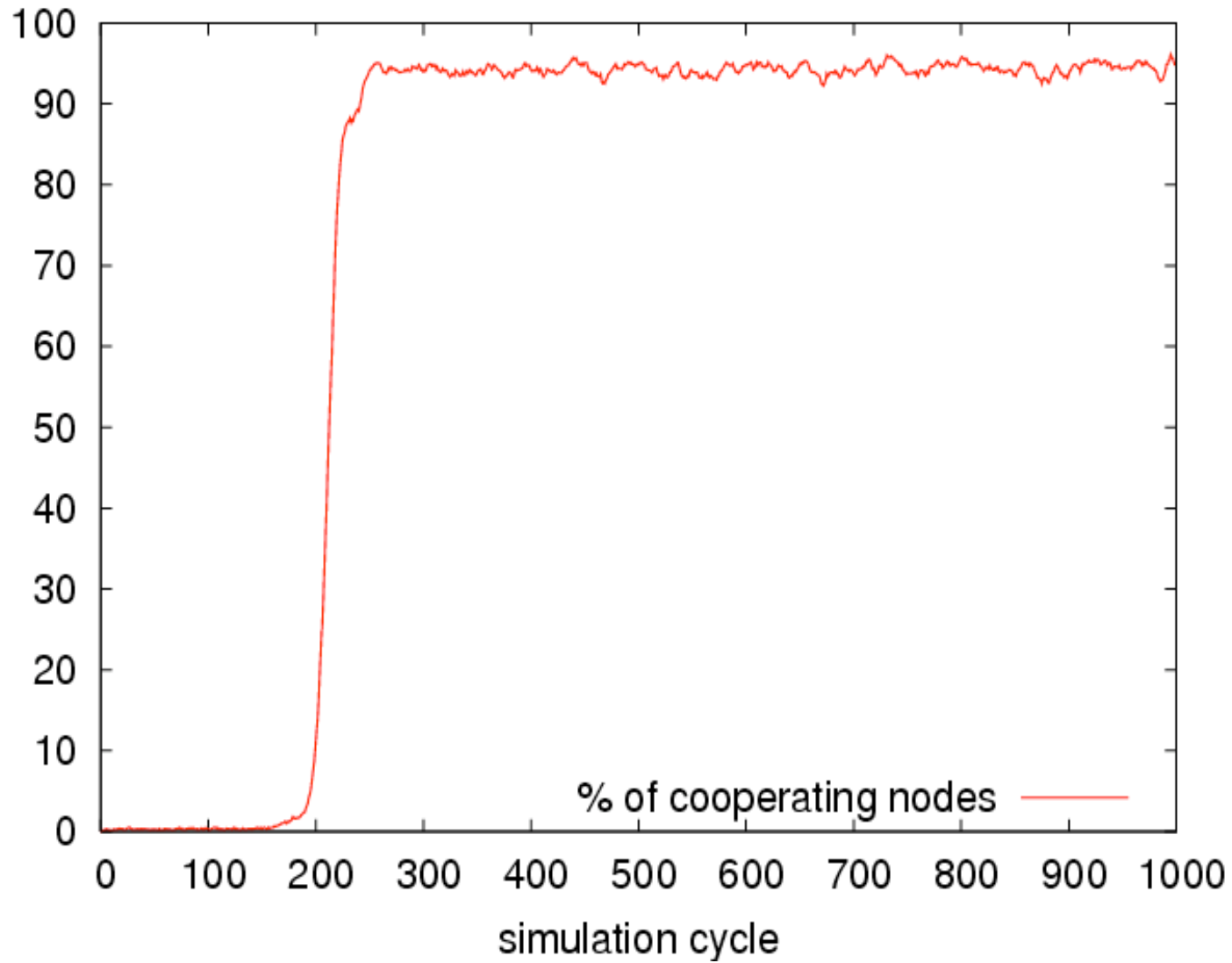# Cycle 220: Cooperation Emerges
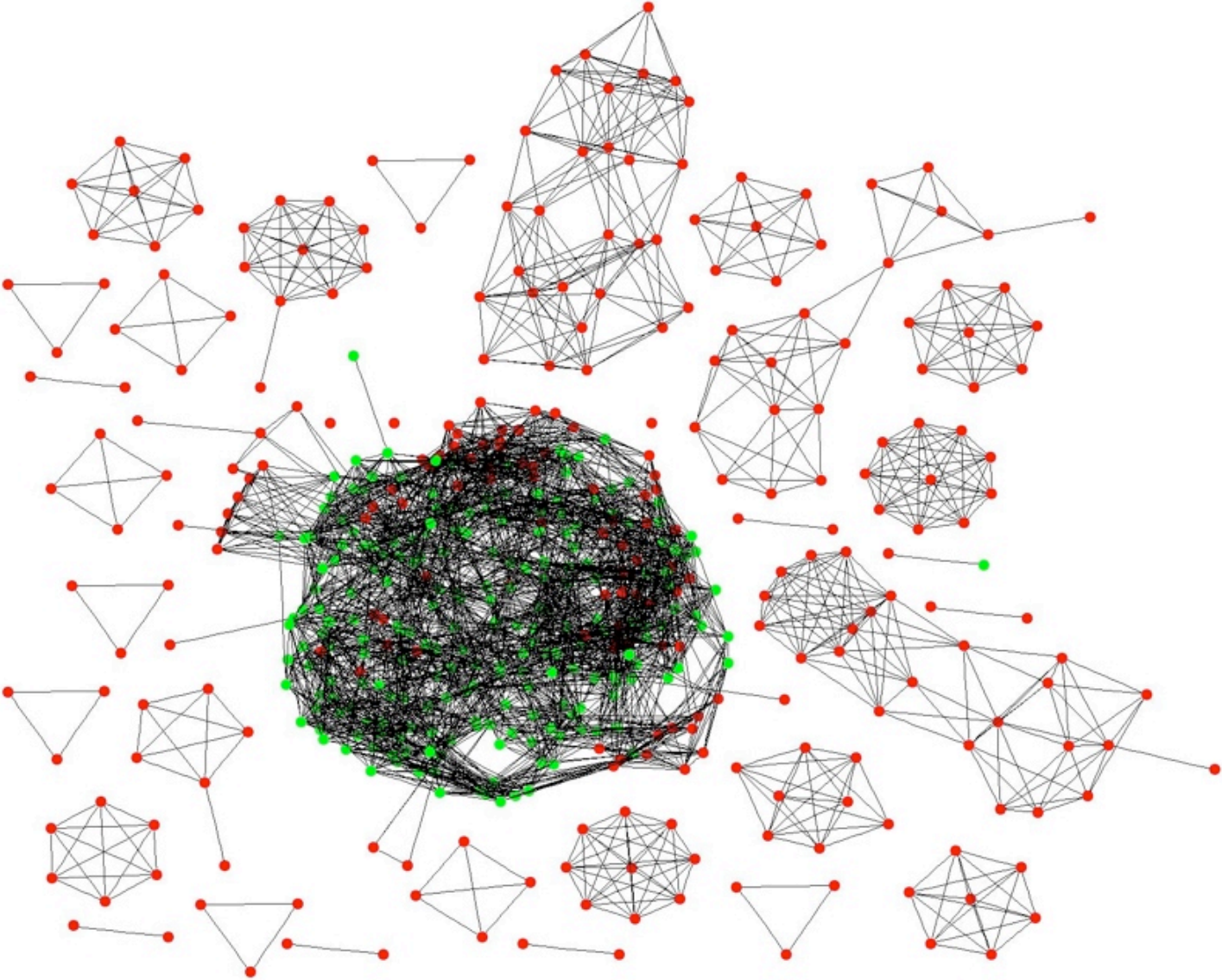
Cycle 230: Coop. Cluster Starts to Break Apart
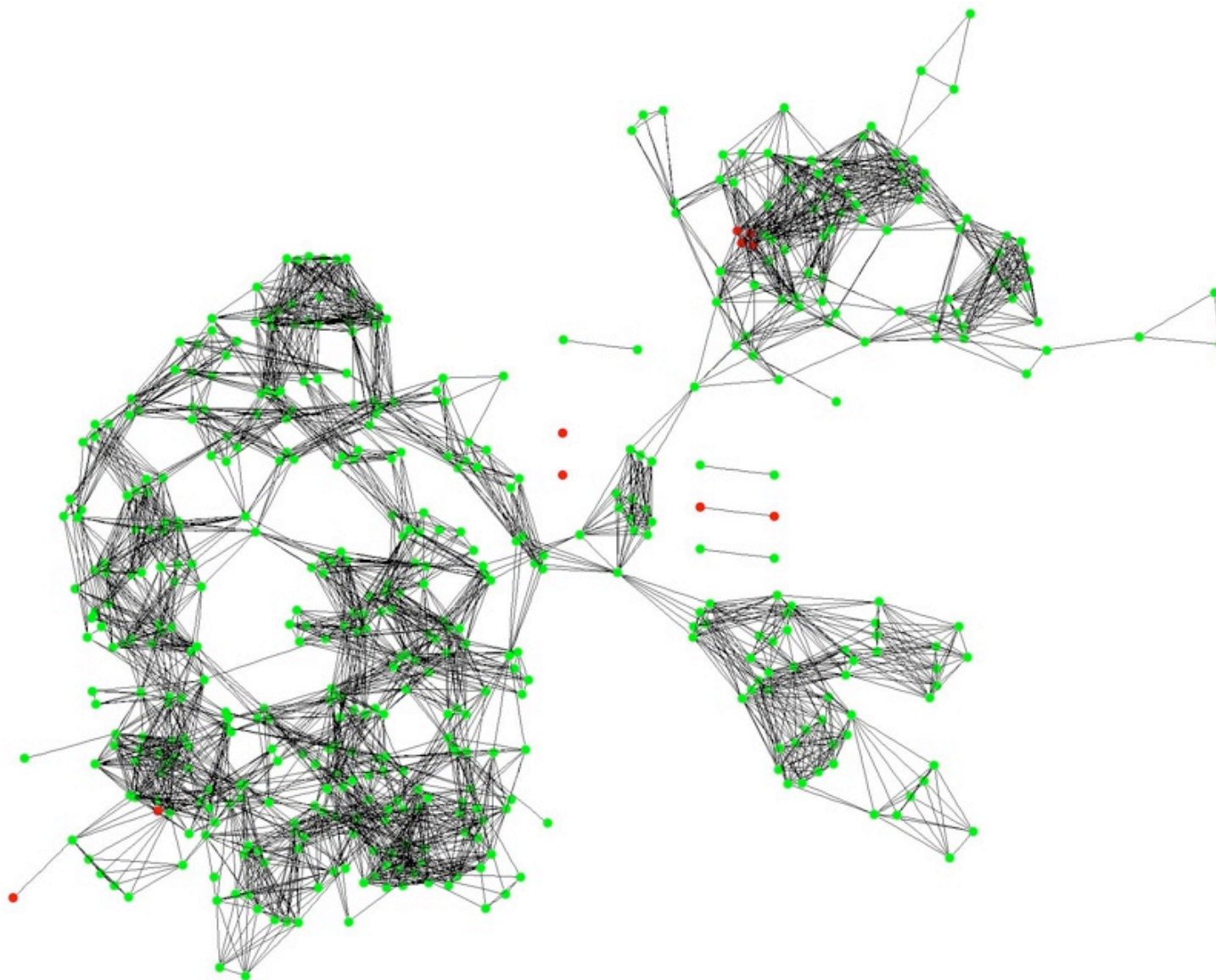
Cycle 300: Defect Nodes Isolated, Small Cooperative Clusters Formed

SLAC Summary

- SLAC produces very high levels of cooperation limits the spread of defection
- Nodes "move" throughout the network to find better neighborhoods
- Group-like selection between clusters
  - Clusters of cooperating nodes grow and persist
  - Defecting nodes tend to become isolated

SLAC and SLACER

- SLAC rewiring mechanism lead to high level of network partitioning

- SLACER: When isolating nodes not all the links are drop. Each link is dropped with given probability W

- Parameter W represents a tradeoff between network randomness and cooperation level

  - W=1: high cooperation, high partitioning

  - W=0.9: high cooperation, small world like topology

  - Low W: low cooperation, random like topology

SLAC and SLACER

W=1   W=0.9   W=0.3


SLAC   SLACER
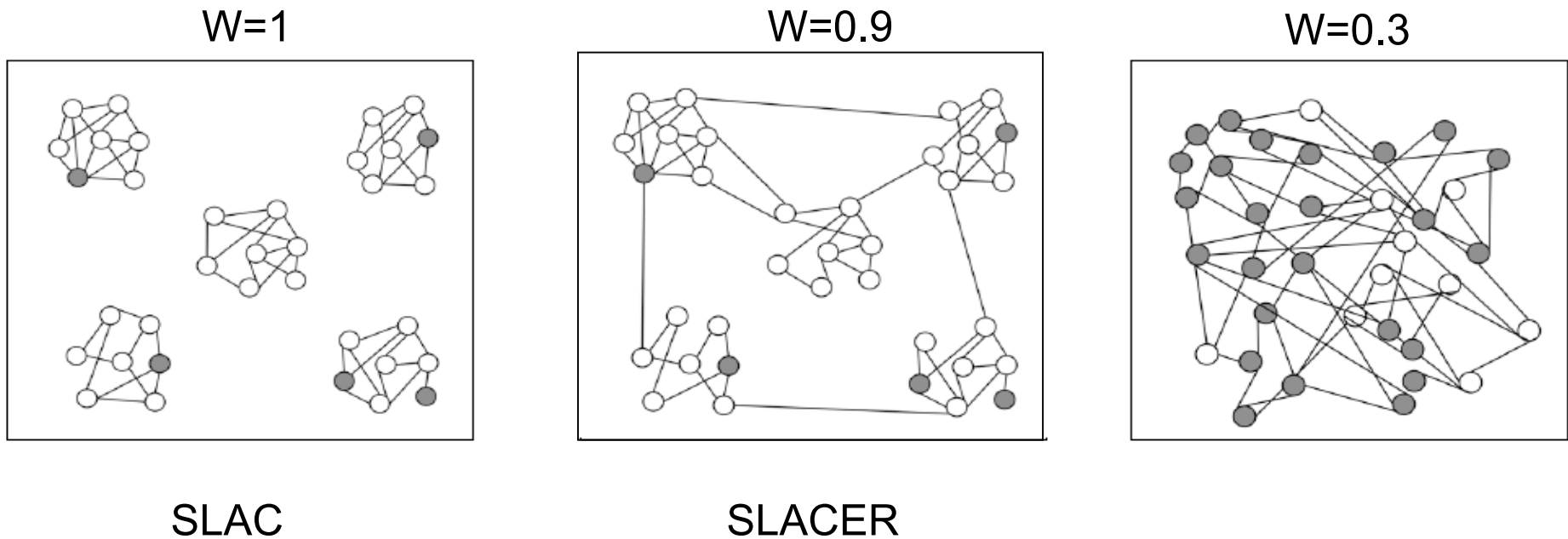
As W is increased (probability of dropping a link when moving) then the network becomes more random and cooperation reduces. Intermeidate points give small-world fully connected networks

SLAC and SLACER

- We applied variants of SLAC and SLACER in P2P applications:

- File-sharing

- Content replication for webservers

- Job sharing requiring specialisation in the clusters in addition to cooperation

Experiments with people
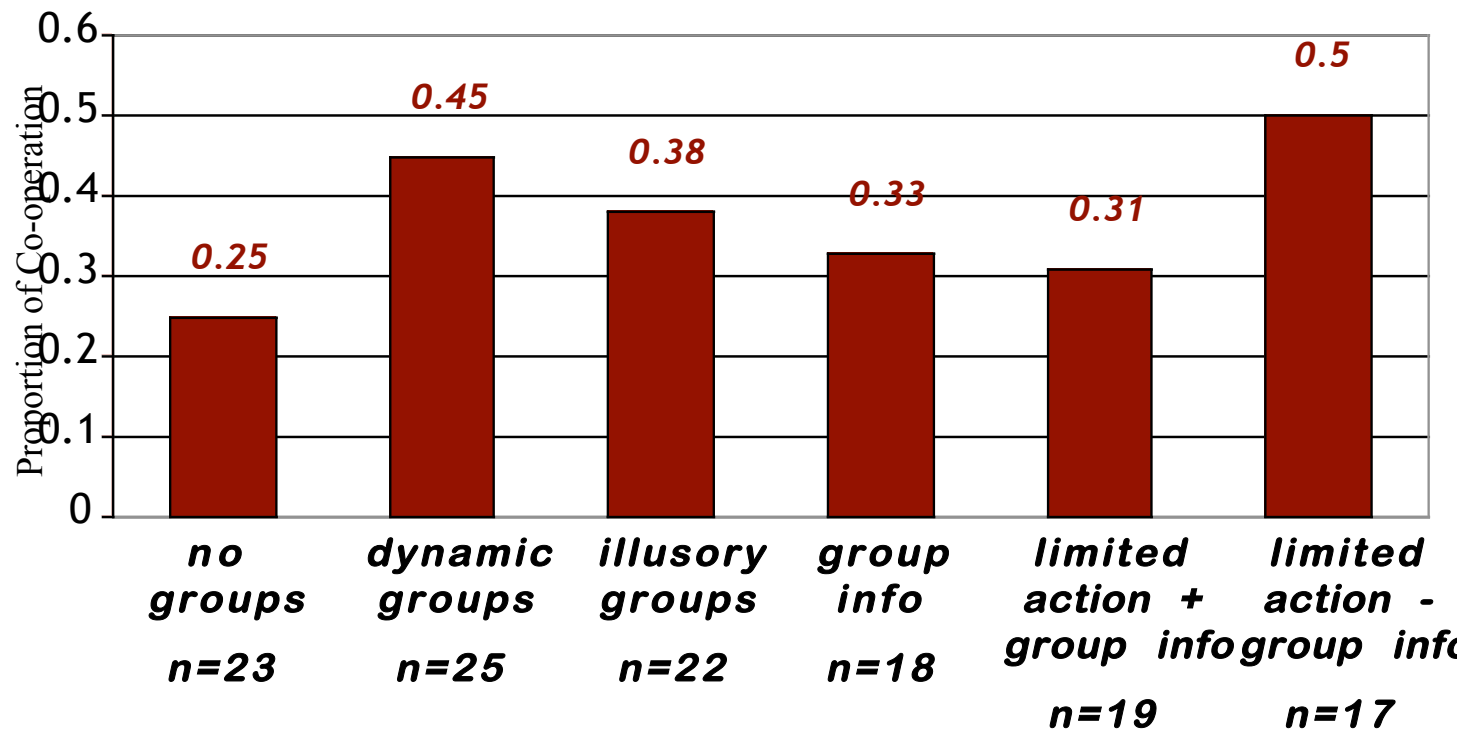
Pilot study. Joint work with Jeremy Goslin, Dept. of Psychology, University of Plymouth (paper forthcoming)

A note on method

- Importing social simulation models into self-* applications is not trivial
- How to do it?
- How we think we did it
  - Start with the abstract model
  - modify in stages towards application
  - Preserve desirable emergent properties at each stage
  - Produce a "chain" of models
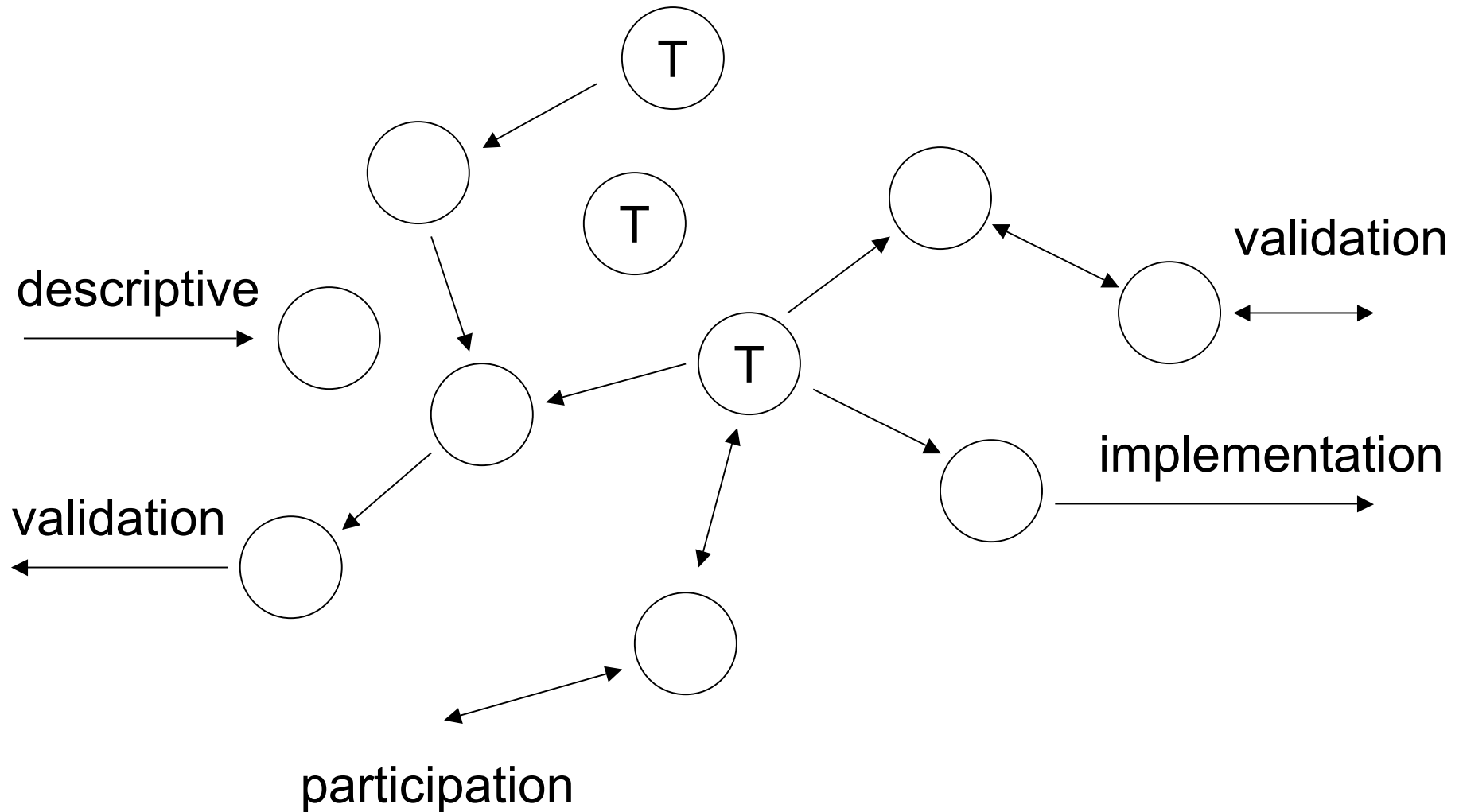
Model chains
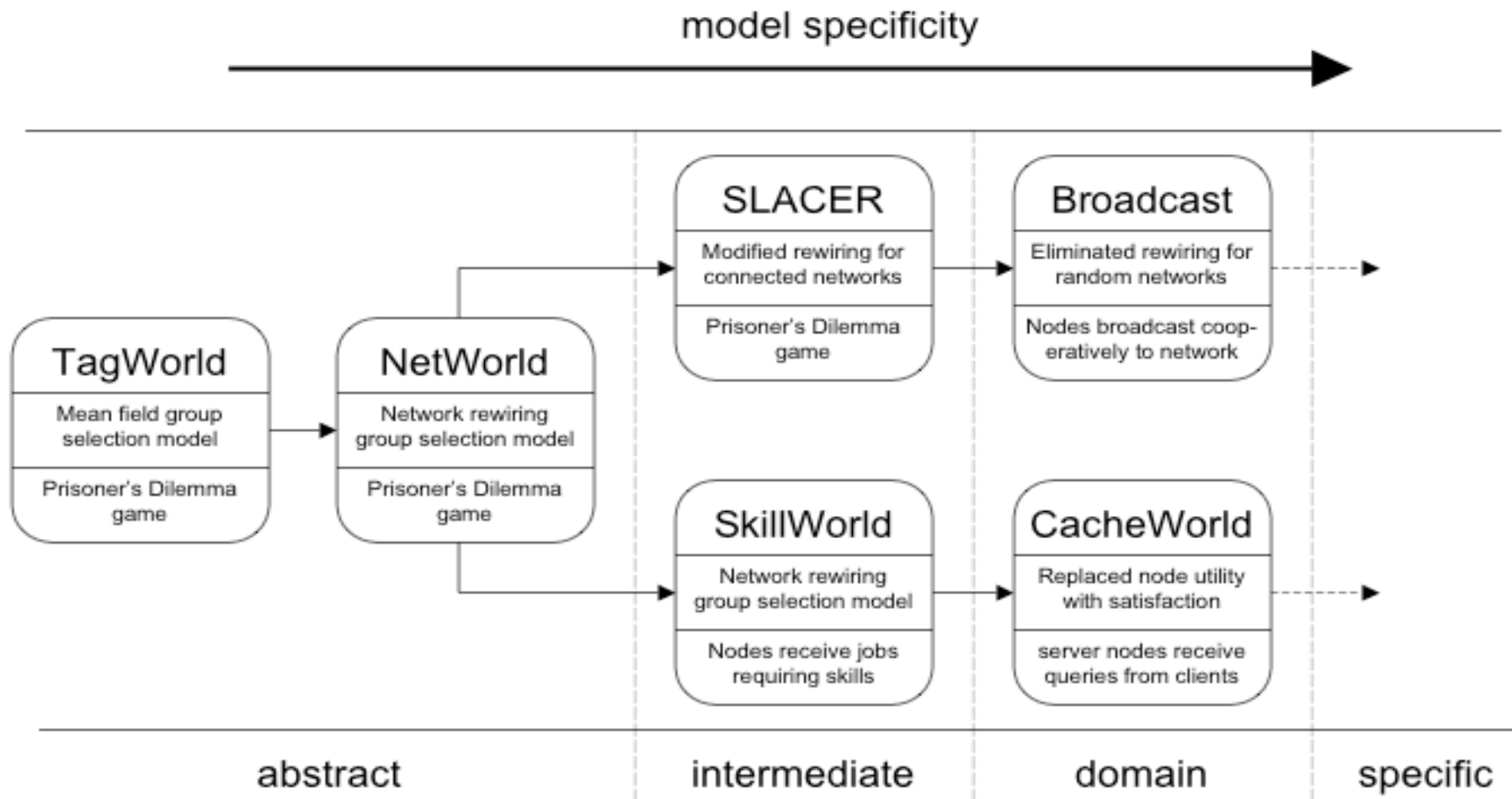
Model chains

- From an engineering perspective "validation" = system works for some application

- However, in social simulation generally, validation = matching / explaining observed phenomena

- Again chains of models can be made from abstract (theory) models to more applied models

Method confusion

- In our community there is diversity of approaches and models
- Theory, abstract, participatory, cross-validated etc.
- This creates confusion and what appear to be endless debates
- But this diversity is a strength!

Model networks – permissive methodology

Finally,

Questions?