Complex Networks and Evolutionary Computation
Carlos Cotta (U. Málaga, Spain)
JJ Merelo (U. Granada, Spain)

Copyright is held by the author(s).
ACM 978-1-59593-698-1/07/0007.

Intro

- What is a network.
- Examples
- Why should we study them

Historical notes

What’s all this about

- Many complex relationships can be expressed by means of networks.
- Networks are composed of nodes (subjects or agents) and edges (directed relations) or arcs (undirected relations).
Graphs
- Directed/undirected
- Geodesics
- Components/cliques
- Degree

Bipartite graphs
- Mapping problem to bipartite graphs
- Some interesting algorithms
- Mapping 2-mode to 1-mode
- Solving the paper assignment problem

Random graphs
- Generation
- How they look like
- Why they are not realistic

Our first social network
- Co-team relations among soccer players.
- Van Nistelrooi, Beckham, Kluivert, Ronaldo and Rivaldo.
In numerical terms

Soccer backoffice conspiracy

- You want to find metastructures in the soccer players relationships.
- Identify core team players.
- Maximize goals score per euro spent.
- Be rich and famous.

Pajek shows

It is obviously connected

- Choose Network ➔ Cohesion ➔ Distance
- Average distance (among reachable pairs) = 1.533
- Distance-based cohesion = 0.389
- It is also reachable.
- Several other things can be computed: geodesics (minimum distance between actors), maximum flow...
But you must obey team discipline

God creates them, and they gather

But classes have always existed

Network links follow a power law

- Log-log links/rank plots show a straight line.
- Sometimes, link abundance plots too.
  - There are 1000 nodes with 1 link, 500 sites with 2 links...
  - Pareto 80/20 rule!
  - Rich get richer!
Which leads to scale-free behavior

- There's no preferred link size
  - Random networks link distribution, as proved by Erdős-Rényi, followed a Poisson distribution.
- Scale-free networks have no preferred scale.
  - Many incoming/outgoing links are unlikely, but possible.

Why do power laws arise?

- Preferential attachment (Barabási)
  - Links are added preferentially to those with links.
- It's not always true
  - Sampling problems.
  - Link decay/aging.
  - Assortative/non-assortative networks.
- Other models: log-normal, stretched exponential

It's a small world

- In small world networks, a few links are enough to connect any two components.
  - Path size scaling with size is logarithmic.
  - Doesn't work in random or regular networks.
- A few links are enough to convert a random network into a small world network.

Complex networks

- clustering
- preferential attachment
- power law
- small world
- giant component
Centrality measures

- Centrality measures indicate the relevance of a node (or link) within the network.
- Measures based on geodesics
  - Closeness
  - Betweenness
- Measures based on degree (or flow)
  - Bonacich power
  - Eigenvector centrality
  - …

Betweenness centrality measures how often a vertex appears on geodesics. High betweenness nodes may control information flow.

\[
C_{i}^{BET} = \sum_{j < k} \frac{g_{jk}}{g_{jk}}
\]

Number of geodesics from node \( j \) to node \( k \) that pass through node \( i \).

Number of geodesics from node \( j \) to node \( k \).
Centrality measures

- Closeness centrality measures how close a node is from the remaining nodes.
- High closeness nodes are the first to get new information (and the most efficient to spread it).

\[ C_{i}^{CLO} = \frac{1}{\sum_{j} d_{ij}} \]

Where \( d_{ij} \) is the length of the geodesic from node \( i \) to node \( k \).

Bonacich power measures the importance of a node's neighbors.
- High power nodes have the ability to influence the network directly or indirectly.

\[ C_{i}^{POW} = \sum_{j} A_{ij} (\alpha + \beta C_{j}^{POW}) \]

Lower than the reciprocal of the largest eigenvalue

Adjacency matrix
Centrality measures

- Eigenvector centrality is a weighted (inversely proportional to length) sum of the number of walks originating at a certain node.

- High power nodes have the ability to influence the network via multiple paths.

\[
C_i^{EG} = \frac{1}{\lambda} \sum_j A_{ij} C_j^{EG}
\]

Reciprocal of the largest eigenvalue

Component ties among New Wave rock bands

Node sizes proportional to eigenvector centrality

Who would win?
(Core/Periphery Concentration Ratio)

Greece (vs Czech); CPCR = 0.816
Czech (vs Greece); CPCR = 1.000

Greece 1 : 0 Czech
C/P Concentration Ratio

Who would win?
(Out-degree Centralization)

Greece (vs Portugal); ODC = 111.81
Portugal (vs Greece); ODC = 262.72
Greece 1 : 0 Portugal
(Final Match)

ODC = 0.00
Each node has 2 out degree

ODC = 0.80
A node has 5 out degree, and all other nodes have 1 out degree

Out Degree Centralization

(\(b = -0.01; \text{BETA} = -0.35\))
Network structure

- A graph partitioning problem
- Number of communities not known
- Two approaches
  - Agglomerative
  - Divisive

Network structure

(Intercommunity links have high betweenness)

Network structure

- Removal of high-betweenness edges results in a dendrogram.

Network structure

- A modularity measure:
  \[ Q = \sum_i \left( e_{ii} - a_i^2 \right) \]

\[ e_{ij} = \frac{\sum_{r \in C_i \cap C_j} A_{rs}}{\sum_s A_{rs}} \]

\[ a_i = \sum_j e_{ij} \]

fraction of edges connecting to community \( i \)

expected value

fraction of edges between communities \( i \) and \( j \)

fraction of edges within a community
Applications in EC - I
- Small world cellular EAs
  - Giacobini, Mike Preuss, and Tomassini
  - Effects of scale-free and small-world topologies on binary coded self-adaptive CEA.
  - Introduces a reproductive restriction, which drastically influences search.

Applications in ECII
- On the Importance of Information Speed in Structured Populations
  - Mike Preuss and Christian Lasarczyk
  - Changing reproduction strategy, small-world networks increase information flow
    - Which might not be good

Applications in EC III
- Evolutionary reconstruction of networks
  - Mads Ipsen, Alexander S. Mikhailov
  - Tries to reconstruct several types of network (including small-world) from its laplacian spectra (set of eigenvalues).

Applications in EC IV
- Small-world optimization algorithm for function optimization
  - Haifeng Du, Xiaodong Wu and Jian Zhuang
  - Kenning: connections is to small world as optimization algorithm is to…
    - Local and long-distance search operators make a small-world network (nodes: solutions; operators: links)
  - Competitive with GAs
Applications in EC V

- Population structure and artificial evolution
  - Arthur Farley
  - Tests different graph structures with mating restriction
  - Small world networks have much the same properties as fully connected networks.

Applications in EC VI

- On properties of genetic operators from a network analytical viewpoint
  - Hiroyuki Funaya and Kazushi Ikeda
  - Studies a GA as a complex network.
    - Populations as nodes, operators as edges.
    - Crossover creates long-distance connections: small world.
  - Worthwhile further investigation

EC coauthorship network

- Scientific coauthorship is a key indicator of the social dynamics of our community.
- The structure of this complex network can provide some insight on the inner workings of EC as a science.
- Interesting questions:
  - How typical a research area is EC?
  - Is the area expanding or shrinking?
  - Are there sociometric stars? If so, who are they?

EC coauthorship network

- Data taken from the DBLP.
  - 7,712 authors
  - 8,501 papers
  - A giant component comprises 62.3% of the network (2nd largest component is 1.4%)
  - Mean distance is 10.9
  - Diameter is 21
  - Clustering coefficient is 0.811
Giant component

EC coauthorship network

Most of the distribution follows a power law

Proceedings editorship

EC follows Lotka's Law of Scientific Productivity

Size of the giant component (percentage)

Year

0 5 10 15 20 25 30 35 40 45 50


Giant component

EC coauthorship network

EC coauthorship network

EC coauthorship network

EC coauthorship network
The community interaction graph shows a core-periphery structure.
EC coauthorship network

- Different centrality measures point to different authors:
  - Betweenness: Goldberg, Deb, Schoenauer, de Garis, ...
  - Closeness: Deb, Michalewicz, Goldberg, Schoenauer, ...
  - Power: Goldberg, Schoenauer, Deb, Keymeulen, ...
  - Eigenvector: Keymeulen, Higuchi, Iwata, Kajitani, …
- Eigenvector centrality prone to hitchhiking.
- Pareto-dominance approach required.

EC coauthorship network

- Resulting non-dominated fronts
  1. K. Deb, D.E. Goldberg
  2. Z. Michalewicz, M. Schoenauer
- Connectedness (not scientific excellence) is measured.

Conclusions

- Complex networks are cool
- And useful

The End

¿Any questions?