NetworkX and igraph

Graph Analysis using Python

Jonas Auel, Sven Fritz

Uni Mannheim

October 28th 2016

Why Python?

- powerful programming language
- allows clear and concise expressions of network algorithms
- growing ecosystem of packages that provide more features
- provides packages in many fields, such as machine learning, statistics and numerics
- in the U.S. Python is by now the most popular programming language for introduction courses

NetworkX

- network creation, manipulation, analyzation (and visualization)
- available for Python
- supported platforms: Linux/Windows/Mac
- load and store networks in standard and nonstandard data formats
- nodes can be "anything" (e.g. images)
- edges can hold arbitrary data (e.g. time series)
- open source

betweenness centrality

- betweenness centrality of a node v: sum of the fraction of all-pairs shortest paths that pass through v
- $$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$
- V: set of nodes,
  $\sigma(s,t)$: number of shortest (s,t)-paths,
  $\sigma(s,t|v)$: number of those paths passing through some node v other than s,t
- if $s = t$, $\sigma(s,t) = 1$, and if $v \in s,t$, $\sigma(s,t|v) = 0$.

igraph

- network creation, manipulation, analyzation and visualization
- available for C/R/Python
- supported platforms: Linux/Windows/Mac
- collection of graph analysis tools
- emphasis on efficiency, portability, ease of use
- open source

hubs and authorities

- authorities: nodes containing valuable content
  hubs: nodes pointing to authorities
- x: vector of authorities
  y: vector of hubs
  A: adjacency matrix of the graph
  α, β: scaling parameters
- $x = \alpha\beta AA^t x$, $y = \alpha\beta A^t Ax$
- x and y are a eigenvector of the largest eigenvalue of AA^t respectively $A^t A$

differences and similarities between the packages

- different syntax (init. graph, adding nodes, drawing graphs, etc.)
- both are very powerful and have many implemented algorithms
- NetworkX has no own drawing tool
- NetworkX only available for Python, igraph is available for R and Python
- documentation on Python is better for NetworkX (igraph has a good R documentation)

references

- Anaconda/Python:
  https://www.continuum.io/downloads
- NetworkX:
  https://networkx.github.io/
- Network data (dolphin and political web blog):
  http://www-personal.umich.edu/~mejn/netdata/
- igraph:
  http://igraph.org/python/
- Dolphin Network paper:
  https://arxiv.org/ftp/q-bio/papers/0403/0403029.pdf