Extracted from:

Complex Network Analysis in Python

Recognize → Construct → Visualize → Analyze → Interpret

Complex Network Analysis in Python

Recognize ➔ Construct ➔ Visualize ➔ Analyze ➔ Interpret

Dmitry Zinoviev
edited by Adaobi Obi Tulton

Complex Network Analysis in Python

Recognize → Construct → Visualize → Analyze → Interpret

Dmitry Zinoviev

Pragmatic Bookshelf

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and The Pragmatic Programmers, LLC was aware of a trademark claim, the designations have been printed in initial capital letters or in all capitals. The Pragmatic Starter Kit, The Pragmatic Programmer, Pragmatic Programming, Pragmatic Bookshelf, PragProg and the linking *g* device are trademarks of The Pragmatic Programmers, LLC.

Every precaution was taken in the preparation of this book. However, the publisher assumes no responsibility for errors or omissions, or for damages that may result from the use of information (including program listings) contained herein.

Our Pragmatic books, screencasts, and audio books can help you and your team create better software and have more fun. Visit us at *https://pragprog.com*.

The team that produced this book includes:

Publisher: Andy Hunt
VP of Operations: Janet Furlow
Managing Editor: Brian MacDonald
Supervising Editor: Jacquelyn Carter
Development Editor: Adaobi Obi Tulton
Indexing: Potomac Indexing, LLC
Copy Editor: Nicole Abramowitz
Layout: Gilson Graphics

For sales, volume licensing, and support, please contact *support@pragprog.com*.

For international rights, please contact *rights@pragprog.com*.

*To my beautiful and most intelligent wife,
Anna, and to our children: graceful ballerina,
Eugenia, and romantic gamer, Roman.*

Understand Egocentric and Sociocentric Networks

Personal social networks are complex networks of persons or social animals. (No, whales and elephants do not have their own Facebook, but if you are intrigued, look at *Animal Social Networks [KJFC15]*!) Respectively, nodes represent people, and edges represent significant social relationships between people: kinship (remember the family tree on page ??), friendship, acquaintanceship, subordination, and the like. Some of the relationships are typically directed (subordination, some subtypes of kinship), and others are undirected (friendship, acquaintanceship), giving rise to the namesake graphs. They may have different weight (*Distinguish Strong and Weak Ties, on page ?*), leading to weighted graphs. One can include any or all of these relationships in one network, ending up with multigraphs and other pseudographs. A social network is truly a complex one!

And the simplest form of a complex social network is an egocentric network.

Egocentric Networks

An egocentric network (or ego network, for short) is the social network of a particular individual. An ego network includes all the individual's contacts and all the relationships among them. Using the terminology from Chapter 5, *Case Study: Constructing a Network of Wikipedia Pages, on page ?*, an ego network includes the nodes from level 1, and the network of subjects related to complex networks is the ego network of the subject "complex networks."

Egocentric networks are used to understand the structure, function, and composition of connections around a single person. Unlike sociocentric networks, they are bounded and focus on individuals (rather than groups).

The central node of an ego network is referred to as *ego* (as in *egoism* and *alter ego*); all the other nodes are called *alter* (as in *alternative* and *alter ego*, again).

To construct a social ego network, start with an ego—say, yourself. Obtain the list of the ego's contacts—the alters. If you explore a social networking website, the list of alters is often called "friends list," "list of subscribers," or "list of followers." You can download it by using the site API, by scraping and parsing the site's HTML code, or, if nothing else works, by copying and pasting the data by hand.

> ## When a Social Network Is Not a Social Network
>
> When your friends say "social network," chances are they are using the words wrong. For example, Facebook is not a social network.
>
> Facebook is a social networking website (SNS)—a website that facilitates social networking by augmenting traditional offline, face-to-face communications with instant online communications. The difference between a social network and a social networking website is like the difference between club members and the club building: while it is easier for the club members to meet in the club building, the building is not strictly necessary for the club to function.
>
> Your mom is still your mom and belongs to your ego network, whether she is on Facebook or not.

If you are a sociologist, anthropologist, or another researcher in the field of social sciences, you may need to deal with real people rather than digital lists. Your principle inquiry tool is probably a name generator (they are described in detail in *Social Network Analysis [KY08]* and other SNA-related books). A name generator is a list of contacts—alters—prepared on your request by the ego person—the person who will be the center of the network. (If you're working on your ego network, you have to make the list yourself.) Often name generators have restricted length to facilitate recollection, but ego networks derived from shortened lists have less elaborate structure.

One way or another, digitally or by hand, you will get a list of some or all alters. You can arrange them into a star network with the ego at the center because all of them are connected to the ego. But that's not enough. Now you have to repeat the contact collection procedure for each of the alters: either by calling the APIs/scraping/copying/pasting or by soliciting names through name generator surveys. With the median number of friends on Facebook being between 155 and 500, depending on whom you trust, the process of data collection may become quite daunting, unless properly automated. One may only wonder how people researched ego networks before Facebook. (Hint: Before Facebook, there was MySpace.[1])

Ego network construction significantly differs from the snowballing process in the way you treat newly discovered nodes. An ego network does not extend beyond the alters. You're supposed to discard any detected node that is not an alter (which is inefficient, but you can save the unwanted nodes for the future—say, for a full social network analysis).

---

1.   myspace.com

You Could Have Had Your Facebook Ego Network

The official Facebook application programming interface (API) v1.0 allowed Facebook apps to download your friends' friend lists. That's all you needed to construct your ego network programmatically. Lada Adamic, a prominent complex network researcher, wrote a program called GetNet that used the Facebook APIs to build ego networks. The program worked well until May 1, 2015, when Facebook retired v1.0 due to privacy concerns. Luckily, I collected my Facebook ego network back in 2013 (see the figure on page 10). It is four years old but better old than nothing.

Once you harvest the ego and all the alters, remove the ego from the network. It is the center of the giant star with the highest connectivity. It dominates the network. It is the tree that makes it hard to see the forest. Removing the ego does not cause any information loss: if needed, just put it back and connect to each existing node.

As an example of a real ego network, let's have a look at my Facebook ego network constructed in 2013. The graph, anonymized for privacy reasons, is shown in the figure on page 10.

Here are some facts about it:

- Each node represents one of my friends, relatives, colleagues or acquaintances.

- Each edge represents what Facebook calls a "friendship," but in reality can stand for anything from "she is my sweetheart" to "what's-his-name."

- The network does not have "my" node because each shown node is implicitly connected to me.

- Some nodes look completely isolated. They are not: each node is connected to me, and there may be other contacts to the nodes that are not my direct contacts.

- Pink and blue nodes represent female and male contacts.

- Some nodes congregate and form dense subgraphs. Some of these subgraphs are all-male, some are almost all-female, and some have mixed gender population. The subgraphs represent different aspects of my social life: family, close friend circle, current and past jobs, and hobbies. An ego network is like a spectroscope that separates your alters into a social spectrum.

- Node size represents betweenness centrality (a measure of node importance; see *Betweenness Centrality*, on page ?).
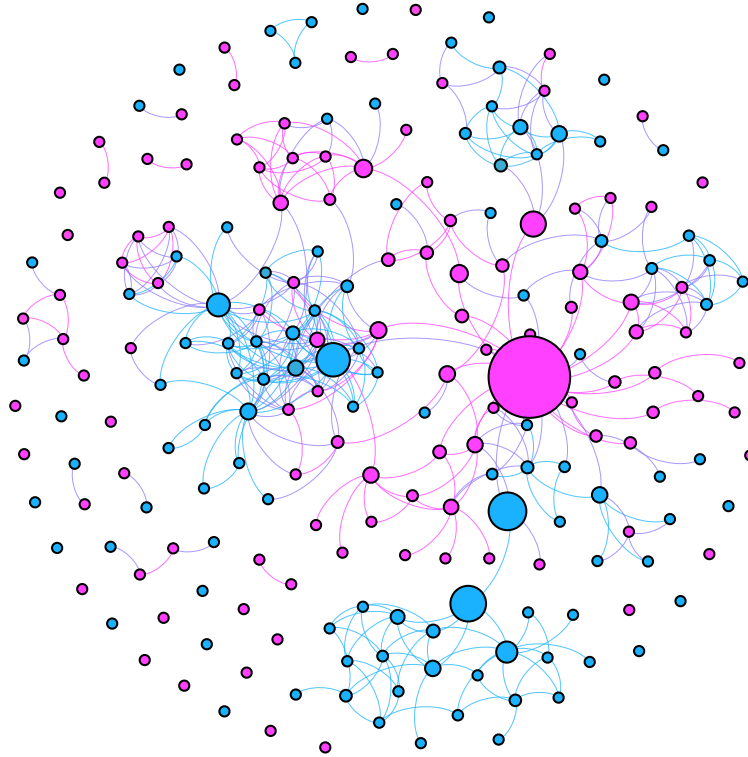
> This fact goes with a tricky question: which node most likely represents my spouse?

Most of these facts are true about most of the human ego networks, but beware: an ego network is only a subgraph of a bigger social network. Anything you measure in an ego network—diameter, centralities, clustering coefficients (Chapter 8, *Measuring Networks*, on page ?)—is an approximation of the same measure for the same node in the bigger graph. Let's next have a look at sociocentric networks, where everyone is an ego and alter at the same time.

Sociocentric Networks

A sociocentric network, or just a social network, is any social network that is not egocentric. Ideally, a sociocentric network is a combination of the ego networks of all egos and includes all relevant (whatever it means to you as a researcher) alters. For example, a social network of all active Facebook users includes all ≈2.01 billion nodes representing active Facebook members (in

2017) and ≈0.25 trillion edges representing their friendships.[2] A complete social network of all living human beings has ≈7.44 billion nodes; the number of edges must be no more than 0.66 trillion if we believe in *Dunbar's number [Dun98]*—150, the number of individuals with whom a person can maintain stable relationships.

A sociocentric network is the prime focus of attention of social network analysts. It reveals all significant relationships of each actor in the network, exposes hierarchical groups of actors, and provides a framework for explaining the structure and evolution of individual edges and node groups.

A non-trivial social network, regardless of its size, is a complex network. What makes it distinctive is not the size but the interpretation: the social theories that stand behind the degree distributions, centralities, local network topology, community structure, and network evolution. The table on page 11 lists some examples of possible social interpretation of complex network properties. Some of them will be covered in this book. In this table, I call nodes "actors" to emphasize their human nature.

Network property	Examples of social interpretation
Local topology	Structural equivalence: if two actors have similar connections to other actors, they are similar or equivalent.
	Triadic closure: two friends of an actor eventually become friends.
	Balance theory: a friend of friend is a friend, a friend of a foe is a foe, and so on.
Degree and eigenvector centrality	Social capital: an actor produces common good for the friends.
	Influence: an actor causes a change in behavior in the friends.
Closeness centrality	Influence: see above.
	Information dissemination/diffusion: how good are actors in broadcasting or sharing information?
Betweenness centrality	Information dissemination: see above.
	Brokerage: how good are actors in serving as "go-betweens"?
Community structure	Homophily (cognitive balance): "birds of a feather flock together."

----

2. bigthink.com/praxis/do-you-have-too-many-facebook-friends

Network property	Examples of social interpretation
	Knowledge preservation: actors in tightly knit communities preserve knowledge.
	Complex contagion: a gang of interconnected infected actors is a source of contagion.
Degree distribution	Small world (six degrees of separation): any two actors on average are connected by six "handshakes."
	Friendship paradox: "my friends have more friends than I do."
Network dynamics	Preferential attachment (Pareto principle): "the [actors] rich [in friends] get richer."

The table is not complete by any means, but it gives you a sample of social research questions and SNA/CNA machinery typically associated with them. If interested, see *Social Network Analysis: A Handbook [Sco00]* and *Exploratory Social Network Analysis with Pajek (Structural Analysis in the Social Sciences) [NMB11]* for a comprehensive coverage of social network analysis with the emphasis on the *social* aspects.