andreas

search

Web Stuff
The Internet
SteamNet
High Stakes
Dot.Crash
SV Work
Networks
HTML Guide
Using SSI
CSS Intro
CSS Detect
CSS Tables
The vi Editor
Characters
PR Release
Animation
Blog
PNG Format
HomeSite
De-Word
Bush Bug
ColorSelect
eCommerce
XML Inside
Site Reviews
Invisible Web
PDAs
eText
Web on PDA
PDFs
Web Obit

Job Stuff
FAQ Stuff
Other Stuff
Personal Stuff
Services

email->andreas

Don't forget to bookmark this page! (CTRL-D)

Linked: The New Science of Networks by Albert-Laszlo Barabasi

Review by Andreas. Contact me at andreas@andreas.com

Albert-Laszlo Barabasi's book on the theory of networks shows that networks (social network of friends, the web's five billion websites, the biological food chain, business and commerce, the growth of cities, intra-cellular proteins, and so on) can be quantified and described with the same type of mathematical laws. These different types of networks share the same properties. By understanding how networks function and grow, one can develop strategies to take advantage of that growth.

Origins of Network Theory

In the 1780s, Euler invented network theory and for most of the last two hundred years, network theory remained a form of abstract mathematics. A network is made up of nodes and links and mathematicians assumed the links between the nodes were randomly distributed. If there are, say, 10 nodes and 50 links, they assumed the distribution would be random and each node would get, on average, five links. Mathematicians explored the properties of these random-distribution networks.

The Social World as a Network

If one applied the model of random distribution in networks to the social world, then six billion humans (the nodes) should each have generally the same number of friends (the links). However, sociologists and economists began to realize that real-world networks were not randomly distributed.

Milgram in the late 60s performed his famous **six-degrees-of-separation** experiment. The popular understanding of Milgram's experiment is that anyone can be linked to anyone else on Earth through only six links. In fact, Milgram discovered:

- **Three Links of Separation**: Some people have such good links that they can get to someone far away with only three links.
- **100 Links of Separation**: Others require up to a hundred links to reach someone else. This also means all of the people within those hundred links were also poorly linked.
- **No Links**: Milgram also found that many people have such poor links that they can't establish a connection to distant others. Many people are isolated into small islands. They are cut off from the rest of society.

In the late 60s, Granovetter, a sociologist now at Stanford, studied how people found jobs. Until then, it was generally assumed that society was homogenous. Granovetter discovered that society is made up of groups of people, which is now known as **clustering**. Granovetter showed that weak contacts were twice as effective (28%) as strong contacts (17%) for finding a job. Casual connections were more likely to lead to a job.

This seems counter-intuitive. It would seem close friends would be better job leads. We tend to gather within groups of similiar interests. If a tennis instructor wants new students, there's no point in asking her friends, who are all tennis instructors. She will find more students by asking people in clusters that have nothing to do with tennis, such as church groups, knitting clubs, and so on. Those clusters (church groups and so on) probably lack tennis instructors. So if you are creating networks, for job hunting, sales, and so on, make lots of casual acquaintances in groups that are outside your normal interests. Better yet, make contacts to the leaders of those clusters, because everyone within those clusters will know the leaders.

(Does this work? That's been my business strategy for the last few years. After the dotcom crash, we diversified our services company **CCG.com** into other industries and regions. We now work with real estate companies, restaurants, and attorneys in other cities, the East Coast, and so on. We work exclusively through personal referrals.)

There's another kind of distribution in social networks. In the early 1900s, Pareto, an Italian economist, discovered the **80/20 Rule**:

- 20% of landowners own 80% of the land.
- 20% of workers do 80% of the work.
- 20% of salespeople make 80% of sales.
- 20% of criminals carry out 80% of crime.
- 20% of websites get 80% of the traffic.
- 20% of the customers create 80% of the calls to techsupport.

The Internet as a Network

The Internet was originally designed to be randomly distributed in order to create a communications network that can survive an attack. In the 90s, physicists began studying the web because it was an example of a network in which all the nodes and links could be tracked. Computer scientists realized that the Web was not randomly distributed. Maps of the web showed that some nodes had huge numbers of links, while most nodes had only a few links.

The Nature of Networks

Barabasi, a physicist, discovered that networks use **logarithmic distribution**, **highly-linked nodes grow faster**, and networks undergo **phase transitions**.

- **Logarithmic Distribution**: Instead of random distribution or bell curve distributions, the distribution of links in a network is determined by logarithmic power laws. If you remember log tables from math, log numbers increase by powers of ten. 2 is ten times larger than 1, 3 is 100 times larger than 1, and so on. This means some nodes have all the links and most nodes only have a few links.
  Earthquakes are measured by log numbers: A magnitude 2.0 is ten times more severe than a magnitude 1.0, a 3.0 is 100 times stronger, and so on.
  On the web, the top websites have ten times more links than the next set, 100 times more links than the third set, and 1,000 times more links than the fourth set. Google's Page Ranking technology is based on log distribution. A website with Google PageRank 5 (PR5) is ten times bigger than a website with PR4, 100x a PR3, 1,000X a PR2, and 10,000X a PR1 website.
  This means the third link at Google is only going to get 1/1,000th the number of visits compared to #1. If you continue down the list, it's extremely unlikely that #25 will get any traffic at all. This works with practically everything on websites: a few pages of a website get most of the visits, most of the searches are based on a few keywords, and so on. They are all based on log number distributions. For example, if you are using Google Adwords for advertising, then you must bid enough to be in the top three positions. Lower than that, you will get very little traffic.
- **Big Nodes Grow Faster**: As new nodes enter the network, they are more likely to link to highly-linked nodes than low-link nodes, because the highly-linked nodes are easier to reach, because they are highly linked. This feedback loop gives preference to the large nodes. Namely, the rich get richer. Networks grow according to the 80/20 rule. Barabasi calls this "preferential linking."
- Networks undergo **phase transition**. This means that when a critical threshold (**the tipping point**) is crossed, the all of the nodes undergo a phase transition and starts acting as a single entity. The property of the network is shared among all nodes in the network. For example, when you boil water, the water acts like ordinary water as it heats up. But at some point, all of the water suddenly starts to boil. There is no "low temperature boiling" or localized boiling. In terms of web markets, there can be a number of dotcoms that are selling dogfood, and at first, the various websites will be different. But when the

market niche crosses a certain size, a few of the dotcoms become very large (the 20%) and the remainder (80%) stay same. But they all take on the properties of the group: they all adapt the same general standards.

Incidentally, this also shows why networks (social, biological, computer, and so on) easily survive most attacks. If a computer virus spreads into a network and destroys perhaps 10% of all nodes, that's not really a problem, because 80% of nodes have low value, so losing many low-value nodes will not affect the network as a whole. However, an attack that targets the large nodes (the 20%) can be catastrophic. The entire network collapses and reverts in a phase transition to an earlier state.

These mathematical laws about networks apply to many kinds of networks: the Internet, wealth and property distribution, membership on corporate boards, personal friendships, intra-cellular protein molecules, and so on.

- Companies that pursue a "business is war" model will be at a self-inflicted disadvantage. They create few links, newcomers don't link to them, business cycle downturns leave them stranded, and so on.
- Companies that embed themselves into the social network of an industry by creating lots of contacts (links) to other companies, suppliers, industry magazines, customers, government, and workers will grow, because the node with the most links will get more links. At some point, the industry (the network) will undergo a phase transition from "just a bunch of separate companies" into an industry. The core companies become institutionalized and they own the industry. Their internal standards become the industry's standard. Pareto's 80/20 Rule applies and the 20% will get 80% of the revenues. Due to the law of preferential linking, newcomers will be effectively locked out of the industry.

One can read the previous paragraph carefully and realize that it applies to many endeavors: international politics, real estate sales, personal networks, and so on.

Barbarasi doesn't seem to know about **geodemographics**. Sociological clustering shows that American society is made up of some 62 clusters. He also does not seem to be aware of the fields of artificial life and concepts of swarming. These fields have developed mathematical models that describe how populations develop and interact.

Links

- Linked: The New Science of Networks. By Albert-Laszlo Barabasi (2002)
- Albert-Laszlo Barabasi at **www.nd.edu/~alb/**
- For another summary of Barabasi's book, see **Douglas Simpson**
- Graph Structure in the Web, by Broder et al. An **IBM research paper**. Fascinating analysis and map of the nature of the web.
- Various articles at **www.nd.edu/~networks**
- Images of networks at **Maps of networks**
- Protein map **Network map of a protein**
- Web map at **Network map of the web**
- Web map at **Network map of the web**

Today is: Friday, 01-Oct-2004 04:15:47 EDT

Page last modified on 07/10/2004

Updates: andreas.com newsletter

I add new pages every month. Sign up with your email and I'll drop you a note (not more than once a month) about new pages. (See more **about the newsletter**.)

Go

home | **web** | **jobs** | **FAQs** | **other** | **personal** | **sitemap** | **legal** |