

Lecture 18: Toy models of human interaction: use and abuse

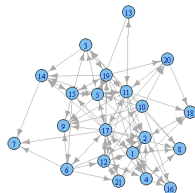
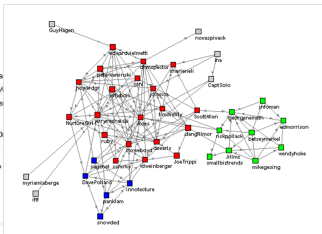
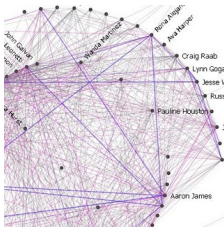
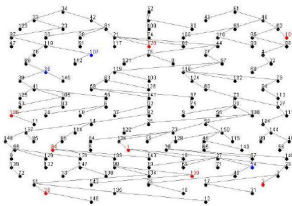
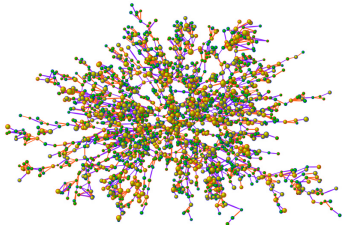
David Aldous

November 2, 2017

Network can refer to many different things. In ordinary language **social network** refers to Facebook-like activities, but mathematicians use it in the context of any kind of relationships between individuals. This has been a hugely popular research field over the last 10-15 years. I will give a quick overview. As the “anchor” for this lecture I will discuss two typical papers, one bad and one better.

What (mathematically) is a social network?

Usually formalized as a *graph*, whose vertices are individual people and where an edge indicates presence of a specified kind of relationship.



In many contexts it would be more natural to allow different strengths of relationship (close friends, friends, acquaintances) and formalize as a *weighted* graph. But the interpretation of *weight* is context-dependent. In some contexts (scientific collaboration; corporate directorships; U.S. Senators) there is a natural quantitative measure, but not so in “friendship”-like contexts. But anyway, instead of always saying “some relationship” let me use the concrete word “friends”.

I won't discuss the **statistical** side of networks – analyzing data. Instead I'll talk about math models. We consider some process in which the individuals interact with their “friends”, and the interaction have some effect – the individual might acquire information, or change an opinion, or gain/lose money, or adopt some new tech item, or catch an illness, or
.....

Mathematically, each individual i has a “state” $X_i(t)$ at time t ; there is some process of interaction between friends, and some “update rule” for how their state changes.

In one sense this is a very general framework. The “social network” graph is arbitrary. The “meeting friends” process might be deterministic (e.g. meet all friends each time unit) or you might meet at random times (this usefully allows for “strength of friendship”). Together these specify a **meeting model**. Then the **update rule** we can choose any state space and any update rule, either deterministic or random.

In another sense we are only considering one-on-one interactions, not broadcast information or majority decisions.

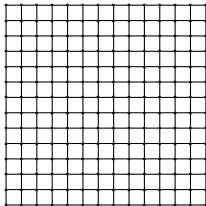
Conceptually the models we use have these two levels, which could be specified independently. But in interesting real-world contexts, usually we don't actually know either

For the social network graph, research articles sometimes use real-world data. But the math theory mostly uses variants of 4 basic idealized networks, for which explicit calculations are comparatively easy.

The 4 popular basic “idealized networks”.

Complete graph: n people, everyone is friends.

d -dimensional grid (discrete torus) \mathbb{Z}_m^d ; $n = m^d$.



Small worlds. Start with the grid, add additional long edges, e.g. chosen at random with chance $\propto (\text{length})^{-\alpha}$.

Random graph with prescribed degree distribution. A popular way to make a random graph model to “fit” observed data is to take the observed degree distribution (d_i) and then define a model interpretable as “an n -vertex graph whose edges are random subject to having degree distribution (d_i) ”. This produces a locally tree-like network – unrealistic but analytically helpful. We will do some math with this **configuration model**.

Here is a common topic, taken from the book **Social Networks and the Diffusion of Economic Behavior** by Matthew Jackson. It refers to “early adopters” of new products. I will quote his verbal discussion, then talk about a math model.

Suppose that we randomly seed the population with some initial adopters of the product. People who have the most friends are the most likely to have friends that are initial adopters. Thus, all else held equal, people with more friends are more likely to then become adopters at a second stage. As we continue to follow the adoption process over time, we see a general propensity for people with greater numbers of friends to adopt earlier on and eventually at higher overall rates (presuming that people with more friends are similar to those with fewer friends along other dimensions).

This prediction is consistent with many empirical studies of diffusion, such as the classic study of the adoption of a new drug by doctors in a famous early study by Coleman, Katz and Menzel (1966).

125 internists, pediatricians, and general practitioners in Illinois were followed over the span of 17 months in 1953-55. Their decisions of when to begin prescribing a new antibiotic, tetracycline, were recorded. After 8 months, those who were in touch with 3 or more doctors who had adopted the drug were more than twice as likely to adopt than those not in touch with any adopting doctor. At the end of the study, after 17 months, those who were in touch with 3 or more doctors who had adopted were approximately 20% more likely to adopt than those who were not in touch with an adopter.

The fact that, all else held equal, more highly connected individuals have higher propensities to adopt suggests a natural comparison between social structures. For instance, if we examine denser social networks that have increased connectedness of individuals, then we expect lower tipping points – how easy it is to get the process started – and higher eventual rates of adoption of the product or behavior. The diffusion has a sort of social multiplier associated with it.

What this means is that once one makes it more likely to have some individuals adopt, it then becomes more likely that their neighbors will also adopt, and so on. In this situation, but slightly increasing the density of social networks, we see a dramatic change in the properties of the system as we leap from one setting where there is no diffusion to another where there is widespread diffusion. As such, it is possible that a relatively small change in the density of social interactions within a population can dramatically affect the end result.

Let's think about inventing a “toy model” to accompany the story above. Take discrete time $t = 0, 1, 2, \dots$. Suppose each person v has a threshold $\xi = \xi(v)$ (possible values $0, 1, 2, \dots$) and will “adopt” when at time $t + 1$ when ξ friends have adopted at time t . So the people with $\xi = 0$ adopt at time 0.

If we specified a graph and the values of the thresholds $\xi(v)$ then we have defined a deterministic process – could simulate. But what about a probability model that we could try to study mathematically?

- Take the $\xi(v)$ to be IID, with specified distribution.
- Take the **configuration model** for the graph, with specified degree distribution $p_i =$ proportion people with i friends.

So each person v has some random number $d(v) \geq 1$ of friends, with distribution (p_i) , and they are “randomly connected” [explain on board].

This produces a locally tree-like network – unrealistic but analytically helpful. I will outline how to set up equations for $q_t(d) =$ probability a person with d friends adopts by time t .

Recall size-biasing and friendship paradox.

Typical person U has $\mathbb{P}(d(U) = i) = p_i$.

But for random friend U^* of U , because of the “random links”

$$\mathbb{P}(d(U^*) = i) = ip_i/\mu, \quad \mu = \mathbb{E}D.$$

This is the **size-bias** concept. Note the setting is slightly different from the **friendship paradox**, which holds for an arbitrary network. The size-bias relation here depends on the particular “configuration model” we are assuming.

We study $q_t(d) =$ probability a person with d friends adopts by time t .

U^* is random friend of U

$$\mathbb{P}(d(U^*) = i) = ip_i/\mu, \quad \mu = \mathbb{E}D.$$

We study $q_t(d)$ = probability a person with d friends adopts by time t .

Key idea:

$$q_{t+1}(d) = \mathbb{P}(\text{Bin}(d, \alpha_t) \geq \xi(U))$$

where α_t is probability U^* has adopted by time t .

[board]

Because U^* has $d(U^*) - 1$ *other* friends,

$$\alpha_t = \mathbb{E}q_t(d(U^*) - 1) = \sum_i \frac{ip_i}{\mu} q_t(i - 1).$$

Now we can numerically calculate the vector q_{t+1} from the vector q_t .

That was a “toy model” – we just “made up” some rules, without any evidence the real world behaves according to those rules, and we did not test against data. Doing this is OK as math, but claiming real-world relevance (other than as possible behavior) is plain “bad science”.

Here is an example of how a toy model paper should be presented,
[show yule.pdf]

Here is a “bad science” example – I show their web page, the actual paper is on my list as *Social consensus through the influence of committed minorities* by J.Xie et al.

[show web page]

Their model is one of many variants of the following **voter model**.

- Each person holds one of two opinions, D or R.
- Each friends-pair meets at random times (rate-1 Poisson process). if different opinions, one (random) person changes opinion.

This is a finite Markov chain, and *eventually* everyone will have the same opinion. The time this takes depends on the graph, but is always a long time (at least $n =$ population size).

Their model changes this update rule, in particular by assuming some percentage of people with one opinion never change opinions.

So this is a ‘toy model’ where rules are simply “made up”; it shows the stated effect is possible, but claiming this explains specific real-world phenomena is bad science.

..... but toy models are fun to study as Math. Here's a model I made up myself. We again recycle the result from the basic toy model for epidemics.

Let's think of a toy model for the spread of epidemics such as influenza. Each infected person will infect some random number of other people; the mean such number is called the **reproduction number** μ . We can use the previous Galton-Watson process to model the number of cases in the initial phase; if $\mu < 1$ the epidemic will not occur; if $\mu > 1$ and there are (at least) several initial cases then there will be an epidemic. Once the epidemic grows it is natural to work with

$$g(t) = \text{proportion of population infected.}$$

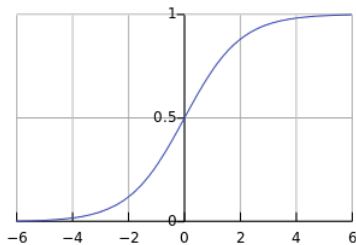
If we ignore the fact that people recover, then the rate of growth of the epidemic is roughly proportion to the number of infected-infected contacts, and this is most simply modeled by the

logistic equation

$$g'(t) = cg(t)(1 - g(t)).$$

The solution, up to an arbitrary time-shift, is

$$g(t) = \frac{1}{1 + e^{-ct}}.$$



Model: Fashionista.

At the times of a rate-1 Poisson process, a new fashion originates with a uniform random person, and is time-stamped. When two people meet, they each adopt the latest (most recent time-stamp) fashion. By general Markov chain theory there is an equilibrium distribution, for the random partition of people into “same fashion”.

For the complete graph network, the analysis is simple. Combining all the fashions appearing after a given time, these behave (essentially) as one infection in the basic epidemic model (over the pandemic window), hence as a random time-shift of the logistic curve F . So when we study the vector $(X_k^n(t), -\infty < k < \infty)$ of proportions of people adopting different fashions k , we expect $n \rightarrow \infty$ limit behavior of the form

