

Social Networks and Mathematical Modeling

E. Ann Stanley

Bend, Oregon

Over the past 30 years, mathematical models of the spread of epidemics and the dynamics of societal cooperation have gradually incorporated more and more information about social network structures. This information is important not just to improve the accuracy of models, but also to improve their qualitative behavior and provide an ever deeper understanding of the spread of diseases and the ways in which individuals interact. Some important questions probably cannot be answered without the use of stochastic simulations that account for the social networks of individual agents.

1. Early Disease Modeling

Mathematical modeling studies have shown that the AIDS epidemic is very sensitive to the human behaviors that spread HIV, including: the amount of risky behavior; the manner in which that risky behavior is distributed in the population; and the social network structures within which people practice those risky behaviors. In fact, these models have shown that if we do not understand all three of these factors, then we cannot hope to predict and control the spread of HIV and other sexually transmitted diseases (Hyman and Stanley, 1988, 1994; Stanley, et al., 1991).

One of the earliest indications that this was so occurred in the mid-1980's. At that time, the US Centers for Disease Control was predicting that the AIDS epidemic was already dying out. They predicted this based upon the fact that the epidemic was growing more slowly than exponentially. With infections that are transmitted via casual contacts, such as measles or the common cold, case data which grow less than exponentially does indeed indicate that an outbreak is peaking and will soon be on the decline. Early mathematical models of the AIDS epidemic also demonstrated exponential growth, followed by slowing towards a peak and then decline (Anderson, et al., 1986)

However, while these early models did account for both sexual activity levels and the distribution of those activity levels in the population at risk, they did not account for social network structure. Instead, they assumed random mixing. Simply adding the fact that people who are high risk tend to associate more often with others of high risk, and similarly for those of low risk, to our model, along

with data on the distribution of partner acquisition rates from various studies in homosexual populations, showed that the epidemic should be growing cubically in time. Reanalysis of the CDC data showed that this was indeed the case (Hyman and Stanley, 1988, Colgate, et al., 1989).

This modeling effort showed that the epidemic was decidedly not dying out. This had important policy implications, since many were saying that, since the epidemic was dying out, it would go away on its own, and it wasn't worth spending a lot of money to control it.

Another important feature of the epidemic that biased mixing models captured was the fact that it was primarily the highest risk individuals who were infected first, followed by the next-highest risk individuals, and so on, whereas the random mixing models had most of their early infections in the large low risk groups.

Modeling HIV spread in age and sex- structured populations pointed out another way in which network structures affect the spread of sexually transmitted diseases. Since female partners tend to be younger than their male partners, women tend to become infected at younger ages than men, with the difference in age being societally-determined, and similar to the difference in age at marriage. But more than that, the way that society is structured, ie the network patterns and norms, strongly influences the speed and pattern of spread. Such questions as who goes to brothels, how many wives and concubines men have, and who gets to have multiple female partners, all affect the spread of disease via sex (Stanley, et al., 1991). Collection and analysis of network data have highlighted some of these factors, showing that clustering of heavily interacting individuals into local risk-taking networks, is an important factor in the spread of disease (Rothenberg et al., 2005 and 2005).

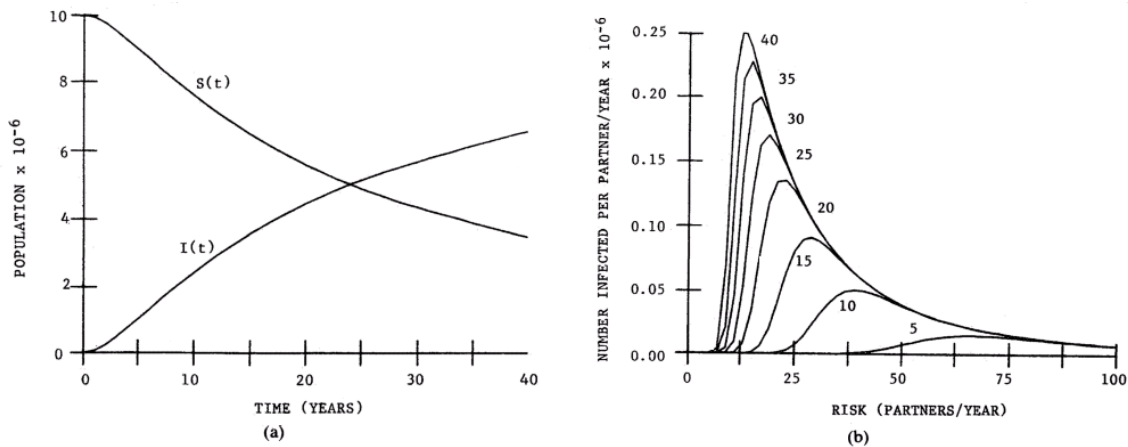


Fig. 1 . The results of a simulation for HIV spread in a homosexual population with biased mixing, where men have a strong tendency to have sex with those who have similar partner-acquisition rates.

(a) Number of susceptibles, $S(t)$, and infecteds, $I(t)$, as a function of time, t . Note that $I(t)$ is proportional to $(t-c)^2$, so that cumulative AIDS cases increase as t^3 .

(b) the distribution of infections with risk at various times. Note the wave that moves from high to low risk. (graphs taken from Hyman and Stanley, 1988)

Many other investigators have looked at mixing patterns as well. Sattenspiel (1987), for example, collected data on Hepatitis A spread among young children in day care centers and devised a mathematical model which used movement matrices to approximate the network behavior. Many researchers have since used various matrices to approximate between group mixing [see Del Valle, et al. (in review) for a brief review], where groups may be divided by age, sex, place of work, domicile and so on.

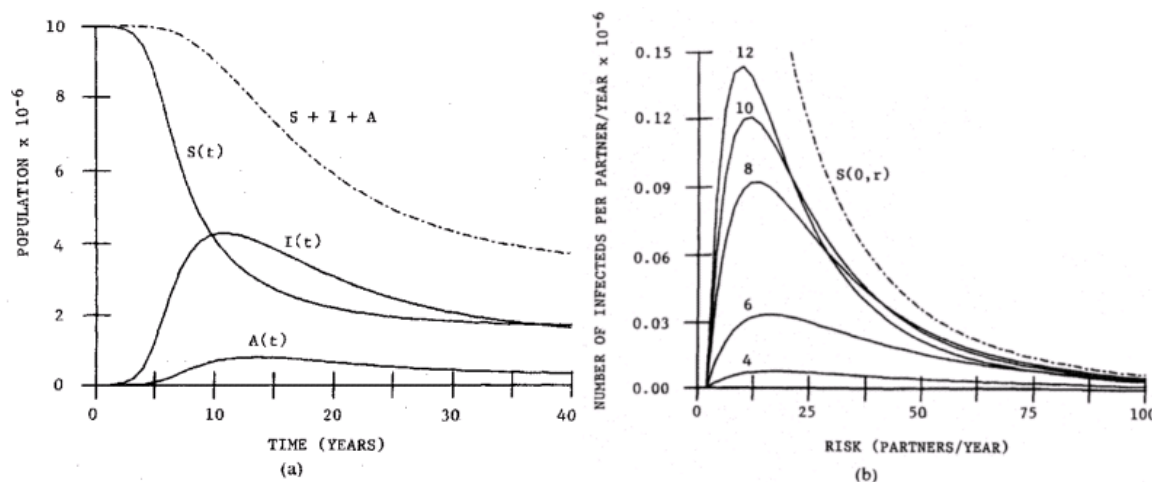


Fig. 2. The results of a simulation for HIV spread in a homosexual population with random mixing and the same parameters as in Fig. 1.

(a) Number of susceptibles, $S(t)$, and infecteds, $I(t)$, as a function of time, t . Note that $I(t)$ increases exponentially at the beginning of the epidemic.

(b) the distribution of infections with risk at various times. Note that the peak only moves slightly in risk with time, and stays at a fairly moderate risk. (graphs taken from Hyman and Stanley, 1988)

2. How do network patterns form, and what types of patterns are likely to emerge under different circumstances?

Beside the question of how networks are actually structured, there is the question of why they have that structure. Why do some people associate with certain other people? How do groups form? What roles do highly connected individuals play?

One way to gain insights into this question is via game theory. Because interactions between people can be categorized as either cooperative or not, it makes sense to use the well-studied Iterated Prisoner's Dilemma (IPD), and then allow individuals to choose and refuse partners based upon the result of previous interactions. The ability to refuse previously defecting partners can allow cooperators to invade defecting populations (Kitcher (1992) and Schuessler (1989)). The ability to seek out known cooperators as partners provides an incentive for agents to be reliably cooperative, so that they will be chosen as partners, or avoid social ostracism (Hirshleifer and Rasmusen, 1989).

One might think, then, that societies would eventually become universally cooperative. But that is not the case. For example, Dugatkin and Wilson (1991) showed that the ability to choose partners in large populations divided into isolated patches may permit roving defectors to move from one patch to the next, taking advantage of the cooperators in the patches. And in a series of papers where we used a genetic algorithm to evolve populations of automata to play IPD with choice and refusal of partners, where choice and refusal was based upon an expected payoff, we found that less than fully cooperative populations could be relatively stable, lasting for many generations (Stanley et al. 1994, Ashlock, et al., 1996, Smucker, et al., 1995), or oscillating back and forth between two different structures.

These less-than-fully-cooperative populations exhibited interesting network structures, sharing similarities with real populations. A few of these networks are shown in figure 3 as undirected "significant play graphs". Note that all simulations were run with 30 individuals.

Analyzing some of these structures showed why they can remain stable, and why there can be a non periodic oscillation back and forth between the “latching” and “Bob and the Raquels” populations, or why two subpopulations can mutually coexist in a symbiotic relationship, as in “stars and hubs”.

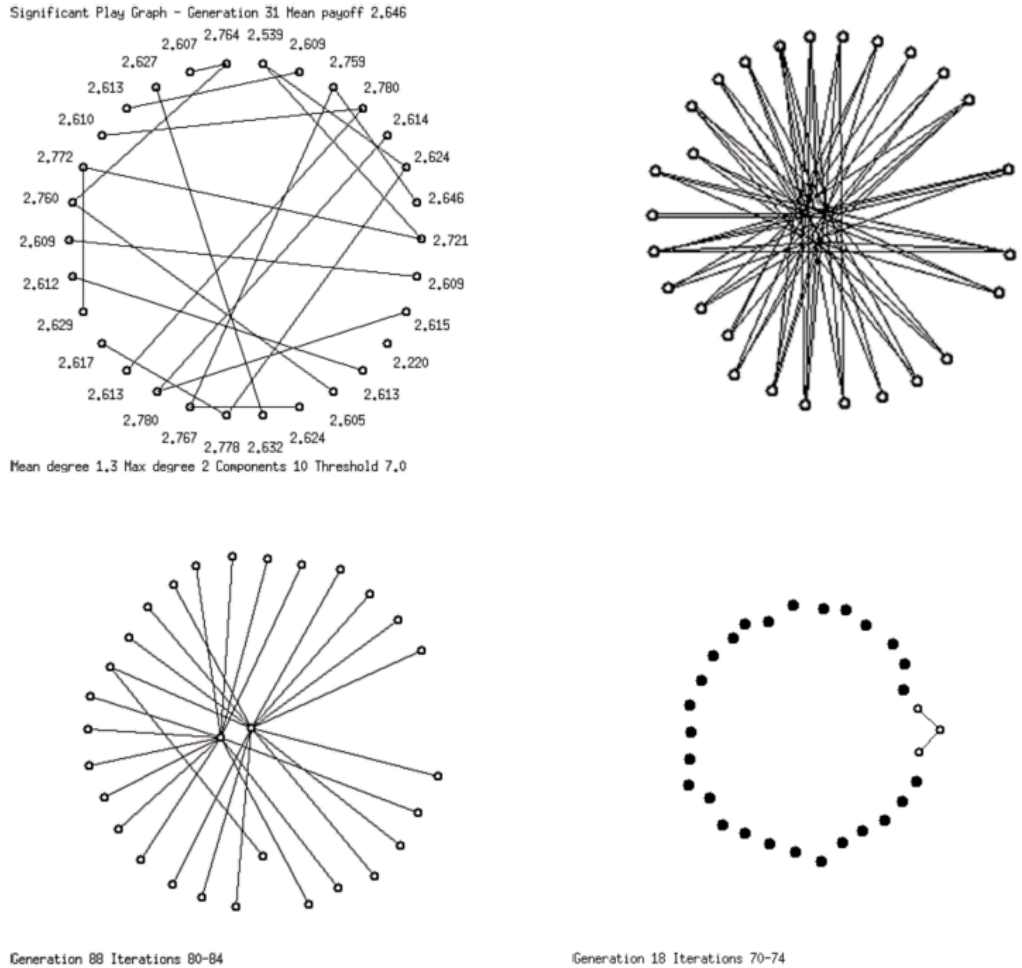


Fig. 3. Different relatively stable networks formed under evolution of IPD with choice and refusal of partners. Shown are the significant play graphs as defined in Smucker, et al. (a) “latchers,” who defect once or twice and then cooperate. (b) “Bob and the Raquels,” where Bobs are “latchers” and “Raquels” are more cooperative than latchers. Raquels occasionally appear, propagate for a few generations until there are too many for the populations to sustain, then disappear. In this case, there are 3 Raquels, who choose each other but don’t refuse the Bobs. (c) “stars and hubs”, which are mutually symbiotic. (d) “Wallflowers” where few individuals play others after the first game (figures from Smucker, et al. and Ashlock et al.).

Of course, this is only a toy model, examined for small populations, so it must not be taken too seriously as an explanation for the appearance of particular network patterns, but it demonstrates how one could develop more realistic models based upon other games with more realistic social constraints.

3. Micro to Macro

In Section 1 we looked at situations where extrapolating macro properties from knowledge of the microstructure of networks led to differential equation models which were able to mimic important

properties of an epidemic and lead to useful insights. However, this is not always the case. Sometimes, a model which contains more details of the network is necessary.

Contact tracing is one of the most promising and controversial methods for controlling the spread of sexually transmitted diseases. In contact tracing infected individuals are asked to name their sexual or needle-sharing partners during a certain period of time, usually for the past six months to a year. Well-trained investigators then track down as many of these partners as possible and convince them to be tested.

Because this methodology has been used effectively for decades with other STDs, many investigators have argued that it should be used to control HIV. Others have argued that it violates people's right to confidentiality, and that people will refuse to come in to be tested if they know they will be asked to name partners. Some specialists in the field argue that the potential for putting people at serious risk of ostracization and physical harm are not worth the potential gain, especially when there is the potential for domestic violence.

Studies of contact tracing as a strategy for controlling HIV, and ensuring that people receive early identification and treatment, have had mixed results, but most have found it to be an effective strategy for finding and counseling infected people. What is certainly the case is that we do not know when it will be effective enough to be worth the risks and when it will not. Mathematical models could help to answer some of the questions surrounding this issue.

We have shown that some aspects of contact tracing can be understood with differential equation models (Hyman, et al., 1994), which have the advantage that model sensitivity can be quickly studied without large numbers of simulations. However, due to the fact that tracing involves finding people who were past partners of those who came in to be tested for one reason or another, such models are of limited validity. In particular, they are not useful when epidemics are spreading rapidly, or when partners are traced for a long time back in the past.

Instead, an approach that directly models the network structure, and is based upon individual agents, such as the model of Kretzschmar, et al. (1996), may be a better approach under these circumstances. Such a model could presumably be expanded to account for snowballing, where partners of partners are aggressively traced. The disadvantage of individual agent-based models is that they cannot easily be analytically analyzed, they require large numbers of parameters, and large numbers of simulations must be run before conclusions can be drawn. Perhaps a mixed approach, such as that of Del Valle, et al. (in review), where mixing matrices are calculated from an individual agent based model, with the intention of using them in a differential equation model, will be the best approach.

While studies of contacting tracing's effectiveness have had mixed results, many have found it to be an effective strategy for finding and counseling infected people and for reaching their communities with appropriately targeted education. The cultural behaviors that tie a particular risk network together (such as the pressure to have an attractive partner) can only be changed in the context of, and by the members of that network. Prevention education within endemic networks that addresses the shared symbolic behavior would likely be more effective than the currently-used behavior change models that deliver more generic and decontextualised information to a wider audience, none of which have produced clear results (Fishbein, et al., 1999).

What is certainly the case is that we do not know when contact tracing will be effective and when it will not. It is also not known how risk behaviors are culturally and symbolically construed in any given network. A network-centric approach, tracing disease as well as culture and concepts, and supported by mathematical models, could help to answer some of the questions surrounding this issue.

With today's computer firepower, more and more researchers are beginning to use simulations of individual agent based models to study the spread of epidemics (see for ex. Newman, 2002; Barthélemy, et al, 2005), and to compare their results with differential equation models.

There are many other situations where the influence of the micro structure of networks, and the individuals within those networks, may need to be modeled with an individual agent-based model, or a cross breed such as that of Del Valle, et al. The spread of technologies, especially in rural populations, and the adoption of new risk-reduction behaviors, can not be understood without first understanding the existing conceptual constructs among key players in the network (MacKeigan and Muth, in press).

CONCLUSIONS

Mathematical models can be used to gain insights into many aspects of the world around us, including the spread of disease and the evolution of culture. These models range from the very micro-models of small populations on networks to the large-scale population models of differential equations. What is clear is that we can no longer learn much of interest from the simple models of the past. Parameters and group structures used in differential equation models need to be informed by network data and analysis, and ideally be extrapolated directly from network data, in order to be most useful. With today's computers, we can begin to link micro structures to macro properties within a single code.

REFERENCES

- Anderson, RM, May, RM, Medley, GF, and Johnson, A. 1986. A preliminary study of the transmission dynamics of HIV, the causative agent of AIDS. *IMA J. Math. Appl. Med. Biol.* 3:229-263.
- Ashlock, D, Smucker MD, Stanley, EA, Tesfatsion, L. 1996. Preferential Partner Selection in an evolutionary study of Prisoner's Dilemma. *Biosystems*, 37:99-125.
- Barthélemy, M, Barrat, A, Pastor-Sattora, R, Vespignani, A. 2005. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *J. Theor. Biol.* 235: 275-288.
- Colgate, SA, Stanley, EA, Hyman, JM, Layne, SP, Qualls, C. 1989. Risk Behavior-based Model of the Cubic Growth of AIDS in the United States. *Proc. Natl. Acad. Sci.*, 86:4793-4797.
- Del Valle, ST, Hyman, JM, Hethcote, HW, and Eubank, SG. in review. *Mixing patterns between age groups using social networks.*
- Dugatkin, LA and Wilson, DS. 1991. Rover: a strategy for exploiting cooperators in a patchy environment. *Am. Nat.* 138:687-701.
- Fishbein M, Higgins, DL, Rietmeijer, C & Wolitski, RJ. 1999. Community-level HIV Intervention in 5 Cities: Final outcome data from the CDC AIDS Community Demonstration Projects. *American Journal of Public Health*, 89(3): 336-345.
- Hirshleifer, D and Rasmusen, E. 1989. Cooperation in a repeated Prisoners' Dilemma with ostracism. *Econ. Behav. Organiz.* 12: 87-106.
- Hyman, JM, and Stanley, EA. 1988. Using Mathematical Models to Understand the AIDS epidemic. *Math. Biosci.*, 90:415-473.
- Hyman, JM, and Stanley, EA. 1994. *A Risk-Based Heterosexual Model for the AIDS Epidemic with Biased Sexual Partner Selection. Modeling the AIDS Epidemic: Planning, Policy, and Prediction*, Raven Press, Ltd.

- Hyman, JM, Li, J, and Stanley, EA. 2003. Modeling the impact of Random Screening and Contact Tracing in Reducing the Spread of HIV. *Math Biosci*, 181: 17-54.
- Kitcher, P. *Evolution of Altruism in repeated oprional games* , Working paper (Dept of Phlosophy, UC San Diego).
- Kretzschmar, M, vanDuynhoven, YTHP, and Severijnen, AJ. 1996. Modeling prevention strategies for gonorrhoea and chlamydia using stochastic network simulations, *AM. J. Epidem.* 144:306-317.
- MacKeigan, T. & Muth, S.Q. (in press). A Grammatical Network of Tzotzil Colour Terms. In *Progress in Colour Studies 1: Language and Culture*, edited by C. P. Biggam and Christian J. Kay. John Benjamins.
- Newman, MEJ. 2002. The spread of epidemic disease on networks. *Phys. Rev. E*, 66.
- Rothenberg, RB, Long, DM, Sterk, CE, Pach, A, Poterat, JJ, Muth, S, Baldwin, JA, and Trotter III, RT. 2000. The Atlanta Urban Networks Study: a blueprint for endemic transmission, *AIDS* 14: 2191-2200.
- Rothenberg, RB, Muth, SQ, Malone, S, Poterat, JJ, Woodhouse, DE. 2005. Social and Geographic Distance in HIV Risk. *Sexually Transmitted Diseases* 32: 506-512.
- Sattenspiel, L. 1987. Population Structure and the Spread of Disease, *Human Biology*, 59.
- Schuessler, R. Exit threats and cooperation under anonymity, *J. Conf. Resolution.* 33: 728-739.
- Smucker, MD, Stanley, EA, and Ashlock, D. 1995. *Analyzing Social Network Structures in the Iterated Prisoner's Dilemma with Choice and Refusal*. University of Wisconsin-Madison Computer Sciences Tech Report CS-TR-94-1259.
- Stanley, EA, Seitz, ST, Way, PO, Johnson, PD and Curry, TF. 1991. The United States Interagency Working Group Approach: The IWG Model for the Heterosexual Spread of HIV and the Demographic Impact of the AIDS Epidemic, The Aids Epidemic and its Demographic Consequences, Proceedings of the United Nations/ World Health Organization Workshop on Modelling the Demographic Impact of the AIDS Epidemic in Pattern II Countries: Progress to Date and Policies of the Future, *UN Publication E.* 91. XIII.5, pp 119-140.
- Stanley, EA, Ashlock, D, Tesfatsion, L. 1994. Iterated Prisoner's Dilemma with choice and refusal of partners. *Artificial Life III, SF Studies in the Sciences of Complexity, Proc.* Vol. XVII, AddisonWesley, pp. 131-175.