# Social Network Analysis With Ashton Verdery

October 10, 2018

Aaron Wagner:

Methodology Minutes is brought to you by the Methodology Center at Penn State.

Aaron Wagner:

Hi, everybody. Welcome to Methodology Minutes, I'm Aaron, your host. With me today is Ashton Verdery, assistant professor of sociology and demography at Penn State. He's also an affiliate of the Population Research Institute, the Institute for Cyber Science, and the Justice Center for Research here at Penn State. Ashton, thanks a lot for sitting down.

Ashton Verdery:

Thank you. Happy to be here.

Aaron Wagner:

Just to ground everybody in what we're talking about, your research is on social network analysis. What are social networks and what is the purpose of analyzing them?

Ashton Verdery:

Social networks are typically conceived as the connections between people and they can be really diversity defined. You can look at connections between people in terms of who had sex with one another, or you can look at connections between people in terms of who just speaks to one another, or more durable connections like who's family members, or who in academic settings, who coauthors with one another. Really, you can define it through sort of any kind of relational event that occurs between sets of people. Social networks as a field is often interested in those kind of individual level or dyadic level connections between people, but also how those aggregate up into broader constellations of connections.

Ashton Verdery:

A really classic example is the small worlds phenomenon, this idea that if I know you and you know someone else, that there's only a few steps between me and other people in that. If you think on a global scale, there's only maybe a few steps between you and a person that is a farmer in Nigeria or between yourself and the President of Japan or something through these acquaintanceship and things like that.

Aaron Wagner:

Yeah, the six degrees of Kevin Bacon phenomenon.

Ashton Verdery:

Exactly.

Aaron Wagner:

Yeah.

Ashton Verdery:

That kind of phenomenon is often replicated through these kind of acquaintances networks, but we see similar kinds of things in other sorts of networks sometimes, for instance, drug injection networks or sexual networks. When you think of a few steps away from you in terms of needle sharing in a drug network or sex in a dating network, you might be able to reach hundreds or thousands of people. So those are properties of social networks that are often interesting to people. I think one of the main reasons that we are motivated to study it is these kind of connections that typically are unseen or at least not traditionally recorded with traditional data collection instruments can have a huge predictive validity or capacity to kind of tell us about what people's behaviors are, and also to kind of tell us about how the broader environment structures like individual decision making and risk. Back to the kind of injection drug network, there's clearly very different risks of like an outbreak of hepatitis C or other kinds of severe infections in a network where people inject with other people, but there's no kind of broader connectivity.

Ashton Verdery:

Whereas in a network where there is that broader connectivity, where maybe a few steps away from you, you can reach 2,000 people or something like that, you're at much higher risk of contracting bloodborne infections and pathogens. Similar ideas can kind of be woven back to sorts of health behaviors and information that we get about, say for instance, migrants, information that migrants would get about opportunities to find jobs in a destination. Those kinds of broader conductivity structures really transmit and constrained sort of the information that's available to us in these important ways. So social network analysis is often about kind of formally studying the connections between people and how they aggregate up into these kinds of broader structures. That could have huge importance for all sorts of activities.

Aaron Wagner:

Thanks. So what brought you to the study of social networks on a personal level?

Ashton Verdery:

On a personal level, I guess there's sort of two ways to answer it. One is I've always sort of been interested in social sciences, like revealing these unseen behaviors or unseen kind of patterns to human activity. The idea that with a few variables, you can predict someone's likelihood of going to college or their income when they're older kind of fascinated me in undergraduate years, and social networks I think are a really nice analogy towards this idea that there are these kinds of unseen patterns that kind of structure our lives. So by measuring and analyzing those things, we can come up with better predictions or better understandings of human behavior, has always been really fascinating to me.

Ashton Verdery:

I guess on a more personal level, I was in college right around when Facebook was booming as a kind of entity. It was only available to college students then, and I remember it was like freshman year or sophomore year, it became available at my university. I thought, "This is so cool," and then you could click through and see like, "Oh, this person that I have a class with knows this other person who's friends with someone else." That sort of small world phenomenon did really just kind of fascinate me on a visceral level that I think connected these more theoretical ideas of these connections between people in such to actual data in a way that you could see the connections that otherwise are very difficult to see and you need special survey instruments. But Facebook kind of made that appealing and visible to any undergraduate in the United States in 2005 or something, or in Canada where I was, so that really connected with me too.

Ashton Verdery:

So I think the twin influence of this, like, seeing the unseen patterns and then having a really staggering example of it as a very popular entity at that time was very interesting to me. I do want to be clear that social network analysis is not just looking at Facebook data. It's a common mistake, but it definitely was an impetus in sort of my interest in the topic.

Aaron Wagner:

Yeah, that's really neat. So you've touched on this a little bit, but other than understanding say the risk of a network of injection drug users, why is social network analysis valuable for people who research behavioral health issues like drug abuse or HIV?

Ashton Verdery:

There are really two principle reasons, in addition to the kind of broader risk network. One, who we interact with is one of the most fundamental parts of our lives, like as social creatures and who we turn to for support and assistance with problems is a really critical part of how we navigate things like drug addiction or rehab and treatment and things like that. So in addition to kind of the bloodborne epidemiological issues, there are real sociological issues about coping with the challenges of drug use, coping with having support to go into rehab and stay in rehab, and then the flip side of it of people that maybe drag you back into the scene or something like that. So there's a lot of reasons to study it from that perspective.

Ashton Verdery:

The other flip side of why it's very useful, it happens to be a very effective way of recruiting individuals into kind of samples from these populations. People who use drugs and people have high risk of HIV tend to be some of the hardest people to survey through kind of traditional methods. Most social science surveys historically have kind of been either someone knocks on your door and says, "Hi, I've got a survey and I'd like you to take it," or they'll mail it to you or they'll phone you at random or something like that, and those are increasingly ineffective methods for reaching general population members.

Ashton Verdery:

But particularly, populations that are highly stigmatized or at high risk of HIV tend to be very reluctant to participate in these kinds of things because their behavior is sometimes illegal in the country where where they're engaged in it, they don't want their membership in that population to be broadly known, or even just from a practical standpoint, "Where do you get a list of all the drug users?" You might have a list of people who attended a treatment clinic, but that's a very select group of drug users versus the kind of broader community.

Ashton Verdery:

So one of the methods that's been found incredibly effective at recruiting these populations over many decades, but particularly since the 2000s is asking individuals who are in the population to refer others that they know that are in the population, and that is fundamentally a social network process. It relies on these social network connections between people, and often because these populations are structured in that sort of small world way where through a few times you can reach a large number of people, it's a very effective way to kind of move through the network and engage large numbers of people.

Ashton Verdery:

The challenge with that is traditional sampling protocols or traditional survey statistics assume that you don't do that, so you need some method to kind of account for that recruitment process versus randomly knocking on doors. So, an additional kind of benefit of social network analysis within studying these groups is it provides some ways and insights to kind of account for that sampling process and create estimates that generalize to the population of interest rather than ignoring that you recruited people that way.

Aaron Wagner:

Thanks. Another population that is hard to reach that you've worked on are migrants, and migrant issues have been much in the news recently as the nation's immigration policies have changed. Does your work suggest anything about the impacts that policies changes have to migrants' ties and migrants' social networks.

Ashton Verdery:

So, my research on migrants focuses on kind of migrant connections to other people, including non-migrants in the destination and people back home, as well as migrants and the destination. One of our main questions is, "Now that migrants can communicate so easily with, across borders and cell phone rates are very low, are migrants maintaining ties to the origin and if they are, what are the ties that they're maintaining?" Going beyond just kind of individual questions of like, "Do people mostly talk to their children or their parents or something like that?"

Ashton Verdery:

There are questions about, "Within a kind of broader migraine community, how is the structure of the community connected through these kinds of ties that people engage in frequently, these communications that people engage with frequently?" One of the big theoretical reasons to care about this is that if the community is broadly connected of people, with the example that we looked at in Mexico, can still access substantial amounts of information from people in the United States over a long period of time through kind of strong network ties. That would suggest that migration to the US could be facilitated for a lot longer than you would otherwise think. That if these ties are really durable and persistent, that the information about job opportunities and things like that in the destination can flow back to the origin in a way that could continue migration kind of like streams or flows. Right?

Ashton Verdery:

So a lot of the recent policies sort of cracking down on aspects of immigration, making migrants more fearful of traveling back and forth and things like that, our research suggests that those policies might affect the strength of ties that connect migrants to their origin communities. Because one of the biggest findings that we have is that migrants who return home, even for kind of brief visits, really keep the network alive and that ties between migrants and family members back home are much stronger and much more durable when migrants return home just for a visit. So to the extent that that might be declining, I think that that would have strong effects on sort of the conductivity that migrants maintain to kind of members of their origin community.

Ashton Verdery:

The flip side of that, on the other hand is that while that is a big impact on sort of the strength of ties between two individuals, we find that the network, even in the absence of that, is relatively robust for a long period of time to kind of this broader connectivity structure between not me and say my brother back home, but between the people I know here and the people my brother knows back home. Those structures, because of kind of redundancies in the network, actually remained pretty robust. So even if the individual ties kind of decline, the overall connectivity structure stays pretty strong, which suggests that there's still a lot of opportunity for information flow and other kinds of flows, remittances and things like that.

Aaron Wagner:

I was actually going to ask you about remittances. Is that something you've studied?

Ashton Verdery:

I haven't looked at remittances as directly as other people, but we've looked at it a little bit and I know a bit about it from reading, but it's not core topic.

Aaron Wagner:

Yeah, I was just wondering if ties would remain strong. I mean, it seems like if you're supporting your mom back in the home country, that those ties would remain strong because you don't want your mom to go without food kind of a thing.

Ashton Verdery:

It's both that, and then there's like a pure coordination thing. The way I always think about is remittances and kind of other things like that are sort of like a second level question. The first level is, "Is the tie still there?" Because if you aren't in contact with someone, how do you send them money? Right? So what we've mostly looked at is sort of these kind of first level contact and connectivity, communication frequency ties, and then I view kind of remittances and other activities like that as being sort of like a subset of those. Because without the contact and conductivity, it's very difficult to coordinate remittances for instance.

Aaron Wagner:

Yeah, sure.

Ashton Verdery:

Okay.

Aaron Wagner:

Thanks. So you mentioned that basically the way that network data is collected means that you're not doing random sampling, and that creates some challenges. What are the largest methodological challenges that, in your opinion, are currently facing social network analysis research?

Ashton Verdery:

I think that there's three. The first is a tension between these sort of snowball or referral-based sampling designs that I mentioned and sort of the random sampling literature. There are social network approaches that use kind of random sampling methods, but what they look at is just an individual and then their ties to, their reports on ties to five people around them. There's no ability to make those kinds of broader conductivity structures.

Ashton Verdery:

The sampling approach that I described, you can make kind of broader connectivity structures with that, sort of like a blended approach. Then there's a third approach where you survey, say everyone in a school. "You ask them, "Who are you tied to?" And then you sort of know, you really know everyone within that school, the conductivity structure. So there are some tensions between those methods, and I think the an exciting area of development are the sampling approaches, which sort of blend the, what's called the saturated network approach where you survey everyone in the school and the egocentric network approach where you do a random sample and ask about five friends for everyone. So, I think that's an exciting area.

Ashton Verdery:

I think one of the biggest methodological challenges is this idea of causality and whether things are moving through networks in a causal way, particularly non infections, but like ideas, information and things like that. Are people's friends and peers changing their behavior or are people selecting their friends and peers on the basis of some behavior that those friends and peers have? Really separating this kind of influence idea that your friends are making you do something, or even subconsciously from like a, you're choosing types of people that accord with the things you are already going to do. This selection idea is one of the really thorny methodological issues for network analysis in general. I think that's probably the hardest one to solve.

Ashton Verdery:

The third big social network challenge right now is how to work with all of the new network data that is available and really how to define what ties are relevant. With location tracking software and things like that, you can get data on, you know, we're in the same room or we were right next to each other, and you could do that for many days then and know who's next to each other and who's not. There's some real challenges with taking the advantages of that big data stuff versus the kind of more traditional social network analysis stuff, which we would go in and ask people, "Who do you spend your time with?" You'd write down five names.

Ashton Verdery:

Really working between those two kind of approaches I think is going to lead to a lot of changes in how we kind of think about a theoretical network issues as well as methodological issues, because they measure different things. Like, who you are near or who you, say email a lot or something like that, these kinds of digitally recorded interactions are meaningful in one way, but in another way they might not be as meaningful because you don't think about it. You know, I probably email a budget person in my office more often than anyone else. But when you think of my academic collaborators, clearly, she's a very vital person to getting my work done, but not in terms of my scientific productivity regarding ideas and things like that. So just looking at email network, you might get a wrong perception of that.

Ashton Verdery:

Similarly, like infrequent contact that is particularly meaningful might be if you asked me who are the important collaborators I have. I might list them out, and some of them I might talk to less often than ones that, if you were to look at my email, that I email with or something like that, but what they say could be particularly weighty and weigh on my mind. So I think there's real challenges with conceptual issues like that that need to be kind of integrated.

Aaron Wagner:

Yeah, those are really exciting. I mean, this is a methodology podcast, so I think it's exciting. I'm not sure what the listeners think, but yeah. Yeah, if you guys could figure out causality, that's pretty-

Ashton Verdery:

Yeah, there are some exciting developments with it, particularly with kind of randomized trial kind of approaches and stuff. There's a lot of very exciting stuff being done in kind of like the more marketing research stuff where they're testing randomized trials of like, limiting people's apps who had shown to and things like that. There was that Facebook study awhile back. We had some ethical issues with it and some methodological issues with measurement, but conceptually, things that are pretty exciting.

Ashton Verdery:

Similarly, there's some actual real, nice statistical advances being made too that are kind of using instrumental variable approaches and things like that are pretty exciting. Another methodological challenge is there've been a number of social network tools that have been developed and measures developed for sort of smaller networks with, say 1,000 to 2,000 people or even fewer, and a certain 1,500 to 4,500 social network ties between the people, and then there's lots of data now, which you can get 60,000 people's email network over a 12 month period.

Ashton Verdery:

So there are real challenges with dealing with time and how to kind of think about networks through time, and then there are challenges with kind of the methodological tools for the smaller network analysis are really appealing and have really great properties, and there's really nice ones for the larger network analysis. But the integration between those two is sometimes missing, and sort of my hobby horse, people, whenever you do anything with social networks, they always want to see the visualizations. Visualizations are like the most exciting thing for everyone who ever sees a network. You see the diagram of the conductivity between two communities and it's very exciting. But often, the visualizations are very challenging or don't work particularly well for the larger networks, and I think that's one of those examples of challenges, that visualizations look great in a small network, but in massive networks it just looks like a yard mall or something.

Aaron Wagner:

Thank you. So your methodological work on network sampling methods has led you to research projects that on some level appear very different in one year studying the heroin crisis here in Pennsylvania and in the other, you are trying to reduce racial disparities in kidney transplants. Can you talk about these two projects and how they pose different challenges in terms of social network analysis?

Ashton Verdery:

Yeah. I think there are people that are sort of dependent variable people and they're interested in outcomes in particular, and then there are sort of independent variable people that are interested in like particular explanations for phenomena that might be diverse, diverse set of outcomes. So, I guess I always thought of myself more as like a explanations person or something, and I think that the social networks kind of approach really is very helpful for understanding both of these phenomenon. Because who we interact with and turn to in times of need and things like that are really fundamentally vital questions that determine both how we access treatment, how we stay in treatment for heroin users, but also who can your transplant candidates go to when they're trying to find a suitable donor to help them or to donate a kidney or be evaluated to donate a kidney.

Ashton Verdery:

So I think that's really kind of the same question in both is like, "How do we access the people we know? How do we make decisions about, of all the possible people could turn to, how do we make these decisions? And what kind of features of the relationship or of the two individuals themselves lead to people seeking out those people versus others and lead to particularly successful versions of those?" In the kidney transplant candidate literature for instance, a very successful strategy, if you just were to look at the data and just ignore the social relations and just say, "All right, people will ask their family members." Who would be the best people to ask? Theoretically, it would probably be their nieces and nephews, because overall nieces and nephews are younger than you, they're healthier than you, and you have usually a lot of them because if you have two siblings and they each have two kids. You've got several of those.

Ashton Verdery:

People don't turn to their nieces and nephews at all. Like, they're very rarely the ones that are asked to help donate a kidney. People are much more likely to turn to their spouse who is the single, in some ways, theoretically, the single worst person you could ask because your spouse is most similar to you in age and health behaviors and is not a genetic match. So there's kind of interesting ways that you can look at these kinds of relations and see like, "Oh, that's sort of a non-optimal strategy, but are there ways that we could reorient it?" This is the purpose of our grant, is to think, "Well, people are most likely to turn to their spouse. Their spouse is often not qualified to donate to them. Are there ways that we can sort of talk to them and figure out ways that they could talk to say other people like their siblings that they have bonds similar to what they have with their spouse?"

Ashton Verdery:

And just like in the heroin network data, we're really looking at who people go to and how their drug use is effective internet works over time, and the results are pretty staggering if you look at it. It's just this real collapse in support for people from when they start using drugs to as they go over time. They lose a lot of the connections to people that don't use drugs and their networks become A, much smaller, much more localized, and primarily other individuals that use drugs and former users. So they sort of lose, they kind of from our results, it looks like sort of entering kind of this world that is very densely saturated with drug users, which I think can be a challenge for seeking treatment and maintaining recovery once you've recovered.

Aaron Wagner:

Yeah. If your whole social network is still using heroin, it's very, very difficult to be out there clean and alone.

Ashton Verdery:

Yeah.

Aaron Wagner:

Yeah.

Ashton Verdery:

This is anecdotally, to me it's roughly knowing, I think alcoholics anonymous and the kind of other organizations have actually, for a long time kind of exploited these ideas without maybe actually explicitly measuring them. But a big part of that is they sort of reorient you into a kind of treatment culture where you're working with a lot of other people who are also rehabbing and dealing with the challenges of that, and you often are kind of not necessarily cutting off, but reducing contact with people that are still using. So there are some kind of anecdotal evidence that kind of successful treatment programs might be using ideas like this. We're just measuring with opioid users.

Aaron Wagner:

That's fascinating. Well, Ashton, really appreciate you taking the time to talk to us today. Thank you very much.

Ashton Verdery:

Yeah, thanks for having me. It was fun to be here.

Aaron Wagner:

You have been listening to Methodology Minutes. Please visit methodology.psu.edu for podcasts, software, reading lists, and more.