Title
Measuring 'neighborhood': Constructing network neighborhoods

Permalink
https://escholarship.org/uc/item/28d4217b

Journal
Social Networks, 34(1)

Authors
Hipp, John R
Faris, Robert W
Boessen, Adam

Publication Date
2012

Peer reviewed

Measuring 'neighborhood': Constructing network neighborhoods

John R. Hipp[1]

Robert W. Faris[2]

Adam Boessen[3]

April 7, 2011

Post-print. Published in Social Networks 34(1): 128-140

Running Head: "Social relation neighborhoods"

[1] Department of Criminology, Law and Society and Department of Sociology, University of California, Irvine. Address correspondence to John R. Hipp, Department of Criminology, Law and Society, University of California, Irvine, 2367 Social Ecology II, Irvine, CA 92697; email: john.hipp@UCI.edu. This research is supported in part by NSF grant BCS-0827027.

[2] Department of Sociology, University of California, Davis.

[3] Department of Criminology, Law and Society, University of California, Irvine.

Biography

John R. Hipp is an Associate Professor in the departments of Criminology, Law and Society, and Sociology, at the University of California Irvine. His research interests focus on how neighborhoods change over time, how that change both affects and is affected by neighborhood crime, and the role networks and institutions play in that change. He approaches these questions using quantitative methods as well as social network analysis. He has published substantive work in such journals as *American Sociological Review, Criminology, American Journal of Public Health, Social Forces, Social Problems, Social Networks, Journal of Research in Crime and Delinquency, Journal of Quantitative Criminology, Mobilization, Health & Place, City & Community, Crime & Delinquency, Urban Studies* and *Journal of Urban Affairs*. He has published methodological work in such journals as *Sociological Methodology*, *Psychological Methods*, and *Structural Equation Modeling*.

Robert W. Faris is Assistant Professor in the department of Sociology at the University of California at Davis. His work focuses on the diffusion of behaviors through networks, with particular emphasis on the ways in which networks of aggression influence, and are influenced by, social structure at the micro level. His work has been published in *The American Sociological Review, Social Networks, Child Development, Complexity,* and the *Journal of Research on Adolescence.*

Adam Boessen is a doctoral student in the Department of Criminology, Law and Society at the University of California - Irvine. His primary research interests include the community of context of crime, spatial analysis, social network analysis, and juvenile delinquency. His work uses quantitative methodologies to examine the relation between residential mobility and crime, the measurement and conceptualization of neighborhoods, and the impact of incarceration on juvenile offenders.

Measuring ‘neighborhood’: Constructing network neighborhoods

Abstract

This study attempts to measure neighborhood boundaries in a novel way by creating *network neighborhoods* based on the density of social ties among adolescents. We create valued matrices based on social ties and physical distance between adolescents in the county. We then perform factor analyses on these valued matrices to detect these network neighborhoods. The resulting network neighborhoods show considerable spatial contiguity. We assess the quality of these aggregations by comparing the degree of agreement among residents assigned to the same network neighborhood when assessing various characteristics of their “neighborhood”, along with traditional definitions of neighborhoods from Census aggregations. Our findings suggest that these network neighborhoods are a valuable approach for “neighborhood” aggregation.

Measuring 'neighborhood': Constructing network neighborhoods

Neighborhoods constitute a fundamental unit of interest for many social scientists. Indeed, sociologists since nearly the dawn of the discipline have focused on neighborhoods as genuine phenomena, as exemplified by the Chicago School in the early part of the 20th Century (Park and Burgess 1921; Shaw and McKay 1942). Psychologists have focused on the effect of neighborhood contexts on numerous individual-level processes (Bronfenbrenner 1977). In the latter part of the 20th Century, the advent of a particular statistical technique—multilevel modeling—coincided with an explosion of interest in the effects of neighborhoods on innumerable individual behaviors of adolescents and adults alike. This "neighborhood effects" literature has looked at the effects of various neighborhood characteristics on delinquent behavior (Osgood and Anderson 2004; Silver and Miller 2004), educational achievement (Ainsworth 2002), low birth weight (Morenoff 2003), and depression (Ross, Reynolds, and Geis 2000), to name just a few. These studies have used numerous conceptualizations of what constitutes a neighborhood, including such varied ecological units as census-defined blocks, block groups, tracts, postal zip codes, or neighborhoods as defined by the cities or residents themselves. All of these studies face a common challenge: how exactly should we conceptualize "neighborhoods"?

A challenge for ecological theories positing such neighborhood effects is that nowhere in this literature is there a very clear definition about what we mean when measuring a "neighborhood". This is not a trivial issue, given that such theories are explicitly ecological, requiring the construction of aggregated measures. Although numerous studies have tested the effects of various neighborhood characteristics on various outcomes, nearly all of these studies are constrained to aggregating the contextual measure of interest to geographic units of analysis that have been designated by the U.S. Census Bureau. There is no reason to assume that these are the only possible neighborhood aggregations that could be constructed. Such arbitrary aggregations can cause the researcher to fail to find an effect that is actually present given a different aggregation of the social context of interest (Hipp 2007).

We propose here one strategy to creating neighborhoods that incorporates information on the social ties within the broader community. Although some adopt a purely geographic conception of the neighborhood, we argue that the presence of social ties is a characteristic of neighborhoods—and are implicit in many existing definitions of neighborhoods—and thus it is reasonable to incorporate the structure of community social ties into a definition of neighborhood boundaries. Although some may believe that our approach conflates the hypothesized positive effects of social ties with the very definition of neighborhood, we point out that our approach: a) considers only the presence of ties and does not presume they have a pro-social character, and b) allows for great variation in tie density. We therefore propose creating a valued sociomatrix of the residents within the community in which the valued relations are some combination of the physical distance between the persons and whether or not they are socially tied. Given the novel nature of our procedure, and the fact that there are myriad options that could be adopted at several of the decision points in our study, we adopt an exploratory approach to demonstrate the utility of this strategy.

In what follows, we first discuss the issues involved in measuring neighborhoods. After discussing why social relations among residents likely play an important role in any definition of neighborhoods, we discuss the several issues that must be addressed when constructing network neighborhoods. We then describe the data and present illustrative examples of our approach, and finally present the results when using various possible configurations of “neighborhoods”. We conclude with a discussion of how our “network neighborhoods” have potential utility for both social network and neighborhood scholars.

Conceptual background

Measuring neighborhoods

The numerous studies that have studied ecological processes in neighborhoods, or the effects of neighborhoods on various individual-level outcomes, contain as a theme the acknowledgement that defining "neighborhoods" is a difficult task. Only occasional studies have offered serious treatments on how we define neighborhoods (Hunter 1974; Schwirian 1983). One theme in this literature, although often implicit, is that neighborhoods are geographic entities that are essentially always defined as contiguous units. Thus, the notion of physical closeness is inherently part of the notion of neighborhood.

The notion of neighborhoods implies the existence of boundaries in the social environment. Being able to define boundaries is necessary when identifying any type of ecological units, ranging from counties to cities to neighborhoods. For ecological units that are political entities—such as cities or counties—the boundaries are usually precisely defined. A challenge for non-political entities such as neighborhoods is that such clear boundaries generally do not exist. The physical environment of cities can appear to be a uniform space in which one street of homes fades into the next. It therefore becomes a challenge for the researcher to identify these boundaries.

Previous research has adopted various boundary definitions. Some researchers have used boundaries defined by the U.S. postal service's zip codes (Harris 2001; Tatlow, Clapp, and Hohman 2000). These are particularly limited given that they were never meant to capture real ecological units, but instead were constructed for the explicit purpose of delivering mail. Other researchers have used units that were designated through some process occurring in the city in which the neighborhoods exist. These "named" neighborhoods often appear to constitute "real" neighborhoods (given that the residents of them often are aware of their names), though researchers have generally not attempted to validate whether they are indeed more effective measures based on some particular set of criteria. Furthermore, some scholars have observed that these named neighborhoods are created through a social process in which the boundaries between certain neighborhoods can be contested. That is, the residents of a lower status neighborhood may claim residence in a higher status adjacent neighborhood; at the same time, the residents of the higher status neighborhood have an interest in defining boundaries that exclude less desirable blocks (Halperin 1998). An analogous process occurs as political parties bicker over the boundaries of districts every ten years as part of the redistricting process. Most commonly, researchers use boundaries defined by the U.S. Census Bureau. Although it is common to lament the use of such officially designated units, it is nonetheless the case that these have the desirable feature of being created with the purpose of constructing something akin to "neighborhoods". Thus, researchers often use block groups or tracts as proxies for neighborhoods (Morenoff 2003; Sampson and Raudenbush 2004; Wooldredge 2002).

When defining the boundaries of neighborhoods, nearly all definitions create boundaries that maximize homogeneity of the residents *within* a neighborhood, and maximize the degree of heterogeneity *across* neighborhoods. This approach is generally taken whether the neighborhoods are named by the residents within the city or whether they are defined by the U.S. Census Bureau. In fact, the Census Bureau adopted an approach explicitly maximizing homogeneity within neighborhoods based on certain key characteristics such as race/ethnicity and socioeconomic status. Named neighborhoods likewise often draw boundaries at points in which the characteristics of residents change. Indeed, a body of research in the geography literature has developed a host of algorithms that create "neighborhoods" based explicitly on the notion of maximizing homogeneity within neighborhoods and maximizing heterogeneity across neighborhoods (for a review, see Duque, Ramos, and Suriñach 2007).

It is useful to ask why nearly all algorithms that create “neighborhoods” attempt to cluster together into neighborhoods residents who are similar on social characteristics. We suggest that this strategy is implicitly based on the presence of social relations, and it is therefore useful to consider the process of tie formation. Although social ties in principle can form between any residents within the larger community, the geography and neighborhoods literature offers two key insights: 1) residents will be more likely to form ties with others who live closer to them in physical space (propinquity); 2) certain social categories create social distance between residents that can also create disjunctions in this structure of social ties (homophily).[1] Indeed, Mayhew and colleagues (Mayhew, McPherson, Rotolo, and Smith-Lovin 1995) suggested conceptualizing a more general concept of distance, with social distance and physical distance as two dimensions of this more general concept. Prior research has documented the tendency to create social ties with others closer in physical space (Butts Forthcoming; Caplow and Forman 1950; Festinger, Schachter, and Back 1950; Hipp and Perrin 2009), and this tendency underlies the notion that neighborhoods will have a geographic component to them.

We suggest that the common strategy in the literature of creating boundaries based on the break points in the social landscape based on the characteristics of residents is done under the implicit assumption that these represent break points in the social relations among residents. Thus, we argue that an implicit assumption underlying this approach is the notion of homophily: residents are more likely to associate with others who are more similar to themselves (McPherson, Smith-Lovin, and Cook 2001). Studies have documented that social distance between individuals can be created by various social categories, including race/ethnicity, economic class, age, marital status, and the presence of children, and this impacts tie formation within neighborhoods even controlling for their physical location (Hipp and Perrin 2009).

[1] There is also some evidence that certain physical boundaries can create disjunctions in the structure of social ties throughout the larger community, although we do not explore this here.

There are several reasons why the pattern of social ties is important for defining neighborhoods. First, social ties can affect residents' perceptions of neighborhood cohesion. Thus, the notion of "neighborhood" often carries with it both a sense of place as well as a community—perhaps partly imagined—of intertwined relationships. The neighborhood literature posits that cohesion is based on residents' perceptions of attachment and similarity of values, and a large body of research has focused on the extent to which residents feel a sense of attachment to the neighborhood.[2] Furthermore, some theories posit that this cohesion can bring about a sense of the neighborhood as a collective unit and can impact various neighborhood-level outcomes such as the level of crime and delinquency (Sampson and Groves 1989; Sampson, Raudenbush, and Earls 1997).

Second, these social ties allow residents to convey information to one another. Residents of neighborhoods can gain information about problems in their neighborhood through these ties. Given the theoretical interest in how residents respond to problems in the neighborhood through various forms of collective action, the flow of information regarding such problems is a necessary precondition for such behavior.

A third mechanism through which social ties might operate is by helping residents provide various forms of collective action, such as informal social control. For instance, social disorganization theory posits that network ties are an important facilitator of residents intervening when they observe delinquent behavior in the neighborhood. For residents to actually engage in collective action behavior in response to such problems requires the flow of information to coordinate such behavior. By providing a sense of cohesion in the neighborhood, social ties can create a sense that others in the neighborhood are also willing to intervene and provide informal social control and sanctioning when observing delinquent behavior in the

[2] Note that this is distinguished from the social network literature, which often measures cohesion with structural measures based on social interaction. The extent to which these structural measures and perceptual measures are related has occasionally been explored in the literature (Hipp and Perrin 2006; Paxton and Moody 2003).

neighborhood. This sense of general willingness to engage in such behavior—what Sampson and colleagues refer to as collective efficacy (Sampson, Raudenbush, and Earls 1997)—arguably rests on the existence of social ties. Indeed, social disorganization studies often posit that such characteristics as residential stability and racial/ethnic homogeneity affect the level of informal social control *because* they give rise to more social ties among residents.

We therefore argue that the notion of network neighborhoods is not really novel, but rather that the idea that networks of relations to define neighborhoods actually underlies many existing approaches, at least implicitly. However, we argue for making explicit that networks are important when creating neighborhood boundaries. Arguably, the fact that few studies have taken such an approach is largely due to the difficulty of collecting such data, and not for theoretical reasons.

Network neighborhoods

An initial question we need to address for constructing network neighborhoods is how to actually measure social ties. That is, what kind of social ties should we focus on when measuring neighborhoods? One approach adopted by Grannis (2009) focused only on neighborly ties. This is, of course, a very narrow definition of the social relationships of persons. An advantage of this approach is that such ties will likely be constrained to very narrow geographic areas. Indeed, Grannis (2009) found that most neighborly ties (defined as the type of informal behavior that occurs with geographic proximity) tended to be constrained to the same block and possibly a few nearby blocks. Of course, this sacrifices quite a bit of information about the total social relations of residents to focus on this one particular type of tie. To assess whether focusing only on neighborly ties is indeed reasonable requires a careful consideration of what exactly it is we wish to measure when accounting for the social relations among residents.

In measuring the social ties of residents, we need to consider what these ties are theoretically expected to accomplish. Given the earlier considerations regarding the roles of social ties, we should determine which social ties will aid in bringing about information flow and a sense of cohesion with the neighborhood. The importance of physical closeness for defining a neighborhood suggests that information flow between persons may not be salient when people live far apart from one another. For example, social ties with persons at a distant work environment may have little meaning for the neighborhood context. Such work ties do not create information flow within the neighborhood, nor do they create a sense of cohesion. It is possible that they may impact the resident's ability to interact with fellow neighborhood residents due to time constraints, though this would arguably be captured by the lack of ties to residents of the neighborhood (Bellair 1997; Hipp and Perrin 2006). Likewise, ties to friends who live further away geographically will have little impact on the neighborhood. However, focusing only on ties to residents on the same block may miss the important ties that link into the broader neighborhood, or even nearby neighborhoods that some have suggested have important effects (Bellair 1997; Gans 1962). Thus, it may be that it is important to focus both on neighborly ties that are extremely localized, as well as ties to nearby areas.

An additional issue to consider is the strength of the ties. For example, should we focus only on strong ties? Or is it important to focus on weak ties, as some neighborhood scholars have suggested (Bellair 1997)? If information flow is of particular interest, it may be desirable to assess the frequency of interaction for ties. If cohesion and attachment are of particular interest, it may be desirable to assess activities of ties that bring about a stronger sense of cohesion. Regardless, it is worth considering whether the valence of the ties should be measured along these various dimensions rather than simply focusing on the presence or absence of ties.

If the researcher indeed has access to all appropriate social ties, then clustering into "neighborhoods" based on social relations is straightforward. However, if locally-based neighborly ties are very common, very geographically constrained (Grannis 2009) and not captured in the measured social relations, then we suggest that one approach to constructing network neighborhoods might use the physical distance between residents as a proxy for neighborly ties, and the actual presence or absence of other social ties to further capture general social relations. Adopting an approach that accounts for both social and physical distance then raises the question of how to relatively weight social and physical distance (Butts and Carley 1999; Butts and Carley 2000). That is, how relatively important are each of these measures of distance? There is no clear answer, as the answer can vary based on the geographic scale, the entity of interest (e.g., people versus neighborhoods) and the characteristics of the ties (e.g., strong versus weak). Little evidence exists regarding this question, as we are aware of only one case study attempting to measure the relative contributions of social and physical distance on neighborhood social ties (Hipp and Perrin 2009).

An obvious challenge is that measuring the location of neighborhoods based on the local density of social ties requires information on all of the social ties among residents within a broader area, such as a city or county. With such information, one in principle could then estimate the boundaries of neighborhoods based on the density of these social ties. That is, we would expect to observe a high density of ties among residents who live within the same neighborhood, and a low density of ties across residents living in different neighborhoods. For example, Figure 1 shows social ties among residents in hypothetical neighborhoods if indeed ties are more likely to form among residents within the same neighborhood.

<<<Figure 1 about here>>>

In contrast, if ties form only based on a particular physical distance function, the pattern of ties across the neighborhoods would not show such discrete breakpoints. Instead, ties would be linked to others closer in space, but there would not be evidence of such clustering.

Furthermore, there would not necessarily be a tendency for ties to be more likely within whatever geographic area is defined as a neighborhood rather than across neighborhoods.

Once the network is defined containing valued ties between the residents of the broader community (based on some combination of physical and social distance), it is then necessary to cluster the egos. There are numerous possible algorithms that can be employed for this problem, and scholars have studied the properties of various clustering approaches and algorithms (for a nice discussion, see Fortunato 2009). It is worth emphasizing that the bulk of these clustering algorithms are designed for dichotomous tie measures (0/1 indicators of the presence or absence of a tie), whereas our approach creates a network of valued relations. These various clustering routines generally yield a solution in which residents are clustered into various groups, or what we consider "neighborhoods". Thus, the boundaries between neighborhoods would be constructed based on the results of these clustering approaches.

A further question when employing such clustering routines for our explicitly spatial problem is whether certain constraints should be placed on possible solutions. Specifically, some have argued that residential blocks are a fundamental unit of geography (Taylor 1997; Taylor, Gottfredson, and Brower 1984), given that social ties are often particularly dense on blocks, and neighborly ties are sometimes constrained to just a single block, at least in urban environments (Grannis 2009). This implies constraining the solution such that households on same block are always considered part of the same neighborhood.

Another question that should be addressed is the often implicit assumption of much prior research that each household must be located in one neighborhood. Must it be the case that a person must live in a neighborhood? Or is it possible to have isolates, and might we consider them to reside in their own neighborhood? In the case of rural areas, it may be reasonable to suppose that geographic isolates indeed live in their own neighborhood. Grouping them into a "neighborhood" with other households that are quite geographically distant arguably does not make conceptual sense. On the other hand, it may not be reasonable to consider those living in more urban areas to be isolates in their own neighborhood. In urban areas, residents are rarely very far away from others. This suggests that the possibility of social ties between a person and other residents is plausible, and that even if such residents choose to socially isolate themselves from nearby residents, there is no reason to suspect that they indeed constitute their own neighborhood. Indeed, such residents are still subject to the same perils in the environment (e.g., crime) that other nearby residents encounter.

Another possibility that is infrequently considered in the neighborhood effects literature is that persons might be considered part of more than one neighborhood. Arguably, a neighborhood is at least in large part constituted by the social and spatial presence of persons. But these persons need not be residents.[3] For example, a person who spends their evenings in and around their home clearly live in the neighborhood surrounding their home, whatever its boundaries might be. To the extent that they talk to their neighbors, they can enhance the information flow in the neighborhood, help foster a sense of cohesion, and provide informal social control by intervening when they observe others engaging in delinquent behavior. But this same person may also spend the daytime several days a week at a workplace that is far removed from their neighborhood. To the extent that their time is exclusively spent within the confines of this work environment while they are at work, they will not have much impact on the neighborhood surrounding their workplace (for example, see Duneier 1999; Jacobs 1961). But to the extent that they are out and about in this area—either walking to and from their car, walking about during their lunch hour, etc, they might be considered part of this neighborhood.

[3] Further examples are persons who regularly are in certain neighborhoods as part of delivery jobs. For example, newspaper delivery, mail carriers, and trash collectors all spend time on a regular basis in some neighborhoods. Of course, it is unlikely that they develop a sense of attachment to the neighborhood, and therefore are more likely to be classified as visitors rather than neighborhood members.

Furthermore, consider a person who frequently chooses to spend their evenings in an area not near their home. This can occur for several reasons: it might occur because they have a particular friend or group of friends who live in this other area; it might occur because this area has amenities that are enticing to the person; it might occur because the area has particular characteristics with which the person identifies. If they only consider themselves to be visitors to this area, this will arguably not be considered their own neighborhood. If they in fact identify with the neighborhood, then this might be considered a second neighborhood for them. This idea has rarely been considered in prior research.

Of course, it is extremely difficult to obtain information on all of the social ties within a larger community. The approach we adopt here uses information on social ties among the adolescents within the schools of a county. Although such data will provide us some interesting insights, they also raise several methodological challenges as we describe below. Nonetheless, we view this as an exploratory study of how network neighborhoods might be constructed. We will compare the results from our approach (giving varying weights to social ties) with Census-defined neighborhood boundaries, as well as a purely spatial approach. We turn to a description of our data next.

Research Design and Methods

Data

The data employed in this study come from the third wave of a multiwave study of a general population sample of adolescents who were identified by school enrollment in three complete school districts in three North Carolina counties and surveyed in school every six months for a total of five assessments. Adolescents in three counties were assessed five times, beginning in Spring 2002 when they were in 6th, 7th, and 8th grades and ending in Fall 2004 when they were in 9th, 10th, and 11th grades. The counties were Moore, Person, and Vance. These counties are rural and located in north central North Carolina. In 2000, the largest cities in these counties were Roxboro in Person County with 8,696 residents, Henderson in Vance County with 16,095, and Southern Pines in Moore County with 11,207. A random sample of the parents of these adolescents was interviewed at three waves, and their assessments of the "neighborhood" are used in the multilevel analyses. Adolescent addresses were obtained from the administrative offices of the school districts at each wave of data collection. Geocodes were assigned to the addresses as the exact latitude and longitude coordinates of a street address by a commercial geocoding firm.[4]

In this study, we consider social ties among adolescents. To obtain adolescent friendship nominations, the data collectors gave each student a Student Directory that included an alphabetical roster of all enrolled students along with a unique four-digit peer identification number for each student.[5] Adolescents used the directory first to identify up to their five closest friends, starting with their best friend. Friends not listed in the directory could be nominated. These friendship nominations comprise the adolescent friendship network.

Assuming that the residents within neighborhoods should have relatively similar assessments regarding the "neighborhood", we use this as a criterion when comparing the quality of the clustering techniques that we utilized. We therefore created several measures of perceptions of characteristics of the neighborhood by combining the responses of parents to the survey. All questions provided 4-point Likert scale responses. For each of these constructs, we combined the scale items with a principal components analysis and constructed factor scores.

[4] Geocoding varies in precision, with the best match being to an exact street address. An exact match is not possible in the cases of misspellings or a rural route or post office box address. When the commercial firm was unable to provide an exact match, the address was re-checked and corrected when possible using U.S. Postal (U.S. Postal Service 2005) and MapQuest (MapQuest 2005) mapping websites. The corrected addresses were then hand geocoded using ArcGIS software or the U.S. Census American FactFinder website (U.S. Census Bureau 2005). If an exact street match could not be made (e.g., because of a postal box address), it was often possible to geocode the address to the ZIP centroid. The average match rate at the block group level was approximately 99%.

[5] For middle-school students, they were given a roster of just those students in the same grade. For high school students, they were given a roster of all other students in the school.

Several measures were based on the responses of parents. For example, we measured *physical disorder* based on three questions about the degree of agreement regarding questions about the neighborhood: 1) people take good care of their homes; 2) your neighborhood is clean; 3) people leave a lot of junk in their yards. We measured *social disorder* based on two questions about the degree of agreement regarding questions about the neighborhood: 1) there is too much drug use in your neighborhood; 2) people respect one another's property. We measure *perception of crime* with one question asking if there is a lot of crime in your neighborhood. *Neighborhood cohesion* is measured with three questions assessing whether people in the neighborhood: 1) are willing to help their neighbors; 2) share the same values; 3) can be trusted. We created a measure of *neighborhood satisfaction* with three questions asking whether the neighborhood: 1) is a good place for your kids to grow up; 2) you would like to live here a long time; 3) is a good place for you to live. We measured *collective efficacy* (the potential for informal social control) based on six questions asking how likely is it your neighbors would step in and do something if teens were: 1) damaging property; 2) showing disrespect to an adult; 3) hanging out and smoking cigarettes; 4) hanging out and drinking; 5) hanging out and smoking marijuana; or 6) a fight broke out in front of someone's house.

We also measure *neighborhood safety* based on the adolescents' responses to four questions about the neighborhood: 1) people feel safe there; 2) people are afraid to come to the neighborhood; 3) people sell illegal drugs there; 4) people there have violent arguments.[6]

Methods

[6] We also created two other measures based on adolescent reports. First, *neighboring* was based on two questions: 1) most of the people know each other; 2) people socialize together there. We created a measure of *informal social control* based on the following questions: 1) adults tell other parents if their child has done something bad; 2) adults keep an eye on what teens are up to; 3) adults would be willing to break up a fight going on there. However, the degree of agreement (based on the ICC's) among adolescents was less than .02 for both of these measures for all of the neighborhood aggregations, suggesting that adolescents do not provide valid sources of information on these neighborhood constructs.

A key focus of this study is attempting to optimally cluster the observed social networks in space. To our knowledge, this has not been done in the literature. We adopt an exploratory approach and use different strategies for creating the valued network of ties between adolescents. We explicitly take into account the physical distance between the dyad members to capture possible local neighborhood ties that are not captured by our friendship tie measure. Given that it is not clear how strongly to relatively weight physical distance and social ties, we take four approaches. In our first approach, we completely ignore the presence of social ties, and just compute the logged physical distance in miles between each dyad.[7] The second approach computes a combined physical/social distance measure by subtracting 1 from the logged physical distance in miles if the dyad is tied. Note that with a logged measure, a one unit decrease is approximately equal to a 63% decrease in the measure; therefore, in this measure a social tie is approximately equal to a 63% reduction in the physical distance between ties.[8] The third approach computed dyad ties by subtracting 2 from the logged physical distance if the dyad is tied: thus a social tie approximately equals an 86.5% reduction in the physical distance between ties. The fourth approach subtracts 3 from the logged physical distance if the dyad is tied, and therefore a social tie is approximately equal to a 95% reduction in the physical distance between ties.[9] Thus, in each case the result is a valued matrix in which each entry is some combination of physical distance between the dyad members and whether or not a tie is present.

As mentioned earlier, there are numerous clustering approaches that can be taken to create estimates of "neighborhoods" based on these valued matrices. We adopted the approach

[7] We log distance rather than keeping it in its original metric given that there is much evidence of a sharp nonlinear spatial decay for social ties (Butts Forthcoming; Festinger, Schachter, and Back 1950; Hipp and Perrin 2009).
[8] For example, a reduction from logged distance of 4 to 3 implies: (exp(3) – exp(4)) / exp(4) = -.632. This ratio is, of course, constant over any chosen logged values.
[9] There are, of course, other possible weighting schemes that could be employed. It is also not clear how to weight social ties that have different characteristics (e.g., strong versus weak ties). Studies might also wish to weight the social ties based on their strength (based on frequency of interaction, emotional attachment, or other criteria). These are all issues we leave to future research.

of using a factor analysis strategy.[10] We then rotated this solution and determined the highest absolute value loading for each individual. The highest factor loading determined "neighborhood" membership for each individual in our first approach to creating network neighborhoods.[11]

As another approach, we relaxed the assumption that individuals must belong to only one neighborhood. In this approach, after performing the factor analysis as described above, we assigned persons to every neighborhood in which they received a factor score greater than 2 in absolute value. We acknowledge that any cutoff value has a degree of arbitrariness to it, and future research would need to assess the sensitivity of the results to such an approach; however, these individuals are two standard deviations away from the mean for such factors, suggesting that it may be reasonable to place them into such "neighborhoods". For individuals who did not have a factor score of at least absolute value 2, we placed them into the neighborhood for which they had the highest value.[12] Note that since the study asked respondents to report on a single neighborhood, it is not reasonable to compare to the other clustering approaches the degree of agreement among members of these multiple network neighborhoods when reporting on the "neighborhood". We will instead simply visually display these multiple neighborhood networks.

[10] Rather than using factor analysis to cluster the observations, another approach devised by Moody (2001) uses the recursive neighborhood mean (RNM) algorithm. This approach 1) creates a series of random variables; 2) computes the weighted mean of all other persons (using the weighted social and physical distance value) adjacent to a node; 3) iterates step 2 a number of times. After a sufficient number of iterations, each person should have values of these initially random variables that are similar to their nearest neighborhoods in social and physical space. Previous work by Moody showed this procedure to show favorable properties on non-valued networks (Moody 2001). However, we tested this approach and did not have satisfactory results using our valued networks. The results we obtained were not at all robust to varying the number of random variables used (which contrasted with Moody's results), and failed to obtain a stable solution even when allowing for very large numbers of iterations (which also contrasted with Moody's results, which obtained a stable solution with relatively small numbers of iterations).

[11] Note that one alternative approach would use the factor weights to create a continuous score of a degree of membership with each neighborhood.

[12] As discussed above, another approach might constrain individuals living on the same block to the same neighborhood. However, this strategy is arguably more sensible when studying a dense urban environment rather than the relatively rural environment studied here, and we therefore leave this to future research. We did create clustered neighborhoods using this strategy, and found generally suboptimal solutions.

An additional challenge is that the school network data do not cover the entire school, but instead are focused on specific grades within each of the schools within the community. Such an approach is reasonable given the high degree of homophily within grade for adolescent friendships in 7^{th}-8^{th} grades, but it does pose an additional challenge for our research question. We adopted the approach of combining the adolescents of a county into a single network.[13] A limitation of this approach is that this full network does not distinguish between ties that do not exist between adolescents in the same school (and therefore could be tied) and ties that do not exist between adolescents in different schools (and therefore could not be tied). Although this is not ideal given that it is possible that a tie in fact exists between the two adolescents, the high degree of homophily within grade suggests that the degree of such error introduced into the study is likely relatively minimal.

We therefore utilized multilevel analysis to assess the degree of agreement (the intra-class correlation) among the residents of our defined neighborhoods on the measures described above assessing the neighborhood. This approach has precedent in the literature to capture the degree of agreement among neighborhood residents as described by Sampson and Raudenbush (1999) in their ecometrics approach. For example, the level one equation for cohesion is:

(1) $$Y_{ik} = \pi_{0k} + e_{ik}$$

where Y is the value of the cohesion scale for individual *i* in neighborhood *k*, π_{0k} is the value of the random intercept for neighborhood *k*, and e_{ik} represents the disturbance for individual *i* in

[13] An alternative approach would perform the clustering on each of the school networks separately, and then attempt to overlay these clustered results into super-neighborhoods. However, this overlaying is not straightforward, and would arguably introduce a considerable amount of error into the procedure. For example, one approach might compute the convex hull around each neighborhood within a particular school network. Doing this for each school network, the degree of spatial overlap could then be computed for neighborhoods across school networks. However, it is unclear how much overlap should be allowable for assigning neighborhoods from different schools or grades to be part of the same super-neighborhood. This arguably introduces an undesirable amount of arbitrariness to the procedure.

neighborhood k. We assume that these errors e_{ik} are independent and normally distributed, with a common variance σ^2.

In the same model we allow π_0 to vary randomly across the k=1 to K neighborhoods, producing the following equation:

(2) $$\pi_{0k} = \gamma_{00} + r_{0k}$$

where γ_{00} is the neighborhood mean and r represents the disturbance in this equation for each neighborhood k which has an assumed normal variance (τ_{00}). We compute the intra-class correlation (ρ) for the estimated models:

(3) $$\rho = \tau_{00} / (\tau_{00} + \sigma^2)$$

Higher ICC levels indicate more agreement among the residents of a particular defined neighborhood. We combined the observations from the three counties into a single dataset for estimating the intra-class correlations.

Results

We begin by describing the clustered network neighborhoods we obtained using our various approaches. Table 1 presents the summary statistics for the number of persons found in each "neighborhood" based on the four clustering methods, as well as the numbers for Census block groups and tracts for comparison. Two features are of particular note: first, the size of the neighborhoods (based on the number of adolescents clustered into each neighborhood) is approximately the size of block groups; second, the size of the average neighborhood decreases as we increase the relative weight of social ties compared to physical distance. Thus, whereas the average block group in this sample has 31.8 adolescents, the average neighborhood based just on physical distance has 38.2 adolescents. As we include information on social ties, the number of "neighborhoods" found based on the clustering method increases (and thus, the average size decreases). Thus, weighting social ties by one yields an average neighborhood size of 37.9

adolescents, weighting social ties by two results in an average neighborhood size of 34.1 adolescents, and the average size is 26.4 when weighting by three. Census tracts are much larger, as the average number of adolescents in them is 88.4.

<<<Table 1 about here>>>

In this same table we also assess the degree of overlap between the neighborhoods detected by each of our clustering methods, and a common U.S. Census geography, block groups. When clustering based just on physical distance, we see that the average amount of overlap between a spatial neighborhood and the block group with which it has the greatest overlap is 62.4%. That is, 62.4% of the adolescents in a neighborhood defined by spatial distance are in the same block group, but 37.6% are in a different block group(s). When taking into account social ties when clustering, and weighting ties at a unitary value, the average overlap is 62.1%, and ranges from 20.6% to 100%. When doubly weighting social ties, the degree of overlap is actually slightly higher (63.3%). However, this overlap falls a bit when trebling the weight of social ties (59.2%). Thus, there is a fair amount of overlap between the "neighborhoods" we detect with our clustering routine and Census defined block groups.

We assessed the probability of ties forming within our network neighborhoods as opposed to across them.[14] As seen in Table 2, ties are more likely to form within network neighborhoods as opposed to across them, and unsurprisingly, this likelihood increases as we increase the importance of social ties. In the model in which a social tie is equated with a one logged unit decrease in physical distance, the odds ratio of a tie forming within network neighborhoods as opposed to across them is 4.9. Even when accounting for the physical distance between adolescents, shared membership in the same network neighborhood increases the odds

[14] We assessed this by estimating simple dyadic models of tie formation. These are logit models in which the outcome is the presence of a tie, and co-membership in the same network neighborhood is the main covariate along with indicator variables for the county of residence. The subsequent models added the logged physical distance variable to the models.

of a tie 43% with the one logged unit weighting, 62% with the two logged unit weighting, and 278% with a three logged unit weighting. In these models, each 1% increase in distance decreases the odds of a tie forming 0.42% controlling for network neighborhood membership when social ties are weighted one logged unit (0.45% when not accounting for the neighborhood). The average distance between these ties is 4.9 miles in Vance County, 6.4 miles in Moore County, and 6.7 miles in Person County.

<<<Table 2 about here>>>

Visual displays of network neighborhoods

To explore this question of overlap a bit more, we next visually display the network neighborhoods discovered here and compare them to standard Census block group boundaries. In all figures, we have jittered the points to preserve anonymity. The figures of the network neighborhoods were plotted in Environmental Systems Research Institute's (ESRI) ArcMap version 9.3.1 with exact latitude and longitude coordinates. Using Tony Palmer's Spider Diagram Tools (http://arcscripts.esri.com/details.asp?dbid=14908), each network neighborhood was plotted by linking each youth with all other youth who are within the same network neighborhood. These "spider" plots use color to distinguish between network neighborhoods. We also plot the 2000 Tiger/Line streets and the census 2000 block group boundaries in all of the maps.

In Figure 2, we plot the adolescents in Vance County along with the network neighborhood to which our approach assigned them (this uses the approach weighting social ties with a value of logged 1). It is satisfying to note that these network neighborhoods tend to be geographically contiguous, given that this is a crucial characteristic defining neighborhoods. Also, there is only minimal overlap between these network neighborhoods. There is a considerable amount of overlap between these neighborhoods and census block groups (the black outlines on the map). Nonetheless, there are numerous points on this map where these network neighborhoods overlap more than one block group, or else are constrained only to a portion of a single block group. Clearly, unique information is obtained with our network neighborhood approach.

<<<Figure 2 about here>>>

Next, we compare the results to two other approaches: one weighting social ties two logged units, and the other allowing membership in more than one "neighborhood." We see in Figure 3 that the resulting network neighborhoods when weighting ties two logged units do not differ greatly from those weighting social ties one logged unit. In Figure 4 we display the network neighborhoods in Vance County when allowing membership in more than one "neighborhood". Unsurprisingly, there is a bit more fuzziness when allowing membership in more than one neighborhood. However, there does appear to be a general pattern in which multiple neighborhood membership for members in the same network neighborhood tends to be to the same one or two other neighborhoods. That is, the second neighborhood membership is not some random connection to another neighborhood in the county, but appears more systematic.

<<<Figures 3 and 4 about here>>>

We provide close-up representations of the downtown portion of Vance County to illustrate these three approaches in a somewhat denser area. Figure 5 illustrates our approach weighting social ties one logged unit, and again shows that some network neighborhoods are constrained to a single block group, whereas others overlap two or three block groups, and one even overlaps with five block groups. In this denser area, there is some evidence of overlap between some of our network neighborhoods. This highlights the effect that the social ties have on these network neighborhoods, as simply using physical distance would not result in such overlap. Relaxing the assumption of strict geography allows a social dimension to these "neighborhoods." The resulting network neighborhoods when weighting social ties two logged units are quite similar, as seen in Figure 6. The multiple neighborhood memberships in Figure 7 do not tend to link members of one neighborhood with varying other neighborhoods, but instead tend to link members of the same neighborhood to just one or two alternative neighborhoods. This may suggest a particular overlap in such neighborhoods, which would provide information that would be useful to model in studies focusing on "contextual effects".

<<<Figures 5, 6, 7 about here>>>

We provide the results for the other two counties in the online Appendix. The general pattern of results is similar in these other counties: importantly, geographic contiguity tends to characterize these network neighborhoods. And the network neighborhoods have only a moderate overlap with block groups, and frequently differ in substantial ways. Another interesting feature we observed across these maps is that many of these ties actually cluster along streets (this can be observed here even though we have jittered the points to preserve anonymity). The fact that streets may actually serve as a *conduit* to social relations is a caution to many approaches that often use streets as *boundaries* for neighborhoods.

Degree of agreement about neighborhood characteristics

We next assess the degree of agreement among adolescents or their parents regarding various "neighborhood" characteristics based on these various network neighborhoods. We display the results of our models in Table 3. For example, the first row shows the ICC's for various defined network neighborhoods when asking parents to assess the amount of crime in the neighborhood: the first column defines neighborhoods based only on spatial distance between adolescents, the second column includes information on social ties as well (with a unitary weight), the third column includes spatial information and social ties (doubly weighted), the fourth column weights social ties by 3, the fifth column combines adolescents based on block groups, and the final column combines adolescents based on tracts.

<<<Table 3 about here>>>

We see in row one assessing the amount of crime in the "neighborhood" that whereas clustering only based on spatial distance, or aggregating based on block group membership, result in relatively high ICC's (.115 and .120 respectively), the highest ICC in fact occurs for our network neighborhood approach which also accounts for social ties (weighted by one), which has an ICC of .125. Thus, the agreement about the amount of crime in the "neighborhood" is highest when aggregating based on membership in this definition of a network neighborhood. We also see that increasing the weight of the social ties actually reduces the quality of the solution based on this criterion: the ICC falls to .096 when doubly weighting social ties, and to .072 when triply weighting social ties. Finally, we see that the ICC is also lower when aggregating to tracts (.103), suggesting that they are arguably too large a unit of aggregation in this relatively rural sample.

We see a similar pattern of results for physical and social disorder. Again, the highest ICC is obtained with our network neighborhoods in which social ties are weighted by one compared to using block groups (.075 versus .056 for physical disorder, and .117 versus .111 for social disorder). Clustering based just on physical distance results in somewhat lower ICC's, suggesting that the degree of agreement regarding the "neighborhood" is improved by incorporating information on social ties. However, increasing the weight of the social ties reduces the quality of the solution. Thus, it appears that a delicate balance between social distance and physical distance is necessary for these clustering results.

Turning to the results for the measures assessing the behavior and attitudes of parents in the "neighborhood", it appears that clustering based on block groups generally does a somewhat better job of producing agreement among residents regarding cohesion, neighborhood satisfaction, or collective efficacy, compared to our network neighborhoods. Thus, the ICC when clustering based on block groups is .117 for cohesion, .135 for neighborhood satisfaction, and .064 for collective efficacy. Again, the ICC's when clustering for tracts is always lower. Although the ICC values are slightly lower for our network neighborhoods, it is still the case that the network neighborhoods using a unitary weight for social ties always do a better job than neighborhoods clustered based only on physical distance, and nearly always do better than network neighborhoods giving larger weights to social ties. This suggests that downweighting propinquity too much is not a desirable strategy when clustering neighborhoods.

Turning to the assessments by the adolescents in the sample regarding the amount of crime and disorder in their neighborhood, the optimal ICC value actually occurs for a network neighborhood with a very high weight for social ties. When weighting social ties by 3, the ICC for this measure is .088, whereas the ICC when aggregating to block groups is .08. The other network neighborhoods do not do that much worse, with ICC's of .075, whereas clustering based on just physical distance results in a worse solution.

Conclusion

This study has explored the usefulness of defining neighborhood boundaries based on the presence of social relations in communities. We proposed an approach creating network neighborhoods based on a combination of the presence of social ties and the physical distance between residents. Given that many neighborhood definitions are implicitly based on the notion of social ties, we have argued that making the connection between the presence of social ties and the formation of neighborhood boundaries explicit is a useful theoretical direction. This allows incorporating information on the fuller network structure to determine the boundary of neighborhoods. An advantage of this approach comes in utilizing information on social ties in a much broader area than just a single census area. However, at the same time a limitation of this approach is the data intensive nature, ideally requiring information on the social ties between all residents within a community. We demonstrated this approach by using data collected on adolescents and parents in the schools of three counties in North Carolina. Although using ties that exist within the context of schools within a county are not ideal, they allowed illustrating the approach. Furthermore, the present study was exploratory given that there are numerous possible choices at each of the several decision points in the process described here.

Our analyses demonstrated that our network neighborhood approach performs satisfactorily. The network neighborhoods we discovered tended to be geographically contiguous, which is a crucial characteristic for the concept of neighborhood. Furthermore, there was a reasonable degree of agreement among the residents in these network neighborhoods when assessing either the crime and disorder characteristics of the neighborhood, or when assessing the relations among residents in the neighborhood. A higher agreement among residents suggests a more homogeneous social environment, which is consistent with the notion of a discrete neighborhood.

We found that an approach that weighted a social tie by one logged unit (a 63% reduction in physical distance) performed best of the approaches studied. An approach ignoring social ties and only using physical distance did not perform as well. And weighting social ties more strongly almost always did not perform as well. In ancillary analyses in which we weighted social ties even stronger (including an approach ignoring physical distance entirely) the clustering solutions either did not converge or yielded too many clusters to be useful. This likely indicates that our social tie information was too sparse to use alone for clustering neighborhoods: this could occur if respondents are asked about too few alters (our study was capturing relatively strong ties) or too few social dimensions. For this reason, our approach of using physical distance to proxy for the presence of neighborly ties--given the evidence of prior research that such ties are geographically extremely local (Grannis 2009)—seemed reasonable.

In an alternative clustering approach, we allowed residents to belong to more than one network neighborhood. Such a strategy has rarely been adopted in prior studies. It is not reasonable to assess the quality of this approach based on comparing the degree of agreement among those in the same network neighborhood when assessing the neighborhood given that the survey instrument for this study constrained respondents to report on a single “neighborhood”. However, it was intriguing to note that when individuals belonged to a second neighborhood there was a general pattern in which this was often systematically related to the specific network neighborhood of membership. That is, adolescents in one neighborhood who belonged to more than one neighborhood tended to belong to the same one or two other neighborhoods. This degree of similarity might imply an important overlap in social context that would be useful to model in future research predicting the effect of context on various behavioral outcomes. Future research will need to assess whether this is also the case in more urban environments: denser areas may enhance the ability to spend time in more than one neighborhood and therefore increase the viability of the multiple neighborhood approach in such environments.

An interesting pattern that we detected is that streets frequently appeared to foster social ties, rather than serve as boundaries. In part, this may be due to bus routes that carry adolescents to schools (sharing a bus route arguably increases the probability of tie formation). Although this may be a unique aspect of focusing on a relatively rural area, it is still the case that streets may increase the probability of tie formation even in urban environments. Given that many approaches defining neighborhoods use streets as boundaries, our finding highlights that defining neighborhood boundaries should be undertaken with caution.

We have acknowledged that this study was necessarily exploratory, and therefore has some limitations. First, the empirical example in a relatively rural area limits the generalizability of the results. Future work should explore this technique in urban areas. Second, as noted at points earlier, more specific information on the form of the social ties would allow for a more satisfactory measure of the social environment, and measuring more dimensions of the social environment (by measuring various relationship contexts, such as friendship, emotional support, casual encounters) would likely increase the power of the approach. Third, other possible clustering techniques may provide even better results—further research is needed to assess whether this is the case.

Fourth, the form of the network questions almost certainly affects the results. Unfortunately, the study design limited the number of network members that could be named to five. This likely dilutes the effect that social ties can have on the solution, and thus increases the effect of physical distance. It is possible that the spatial footprint of these neighborhoods could be larger if more network ties were elicited, though this is speculative. Likewise, the fact that we are limited to ties among adolescents is a limitation to capturing the entire social environment. An additional issue is the fact that adolescents were limited to naming ties in the same school or even grade (for middle schools). This also likely affects the spatial footprint of our network neighborhoods in ways that are hard to assess. Our strategy coded tie information to zero for persons in different schools but the same county. If one thought that such ties across schools exist to a non-trivial degree, an alternative approach would code these ties as missing and use an imputation scheme. This might be useful to assess in future research.

In conclusion, we have suggested that network neighborhoods are a useful direction for neighborhood scholars. We have emphasized that nearly all existing conceptualizations of “neighborhood” are based on the presence of social relations, and that many existing strategies for defining neighborhood boundaries are implicitly predicated on the notion of homophily. Explicitly incorporating information on the presence of social ties will allow the neighborhood effects literature to create more socially meaningful neighborhood boundaries. This approach takes into account the fuller network structure to assess both the boundary of neighborhoods, as well as the possible consequences such social ties can have for the residents of neighborhoods.

References

Ainsworth, James W. 2002. "Why Does It Take a Village? The Mediation of Neighborhood Effects on Educational Achievement." *Social Forces*:2002.

Bellair, Paul E. 1997. "Social Interaction and Community Crime: Examining the Importance of Neighbor Networks." *Criminology* 35:677-703.

Bronfenbrenner, Urie. 1977. "Toward an Experimental Ecology of Human Development." *American Psychologist* 32:513-531.

Butts, Carter and Kathleen Carley. 1999. "Spatial Models of Network Formation." *Working paper*:1-36.

—. 2000. "Spatial Models of Large-Scale Interpersonal Networks." *Working paper*:1-47.

Butts, Carter T. Forthcoming. *Space and Structure: Models and Methods for Large-Scale Interpersonal Networks*, Edited by S. E. Fienberg and W. J. Linden. New York: Springer.

Caplow, Theodore and Robert Forman. 1950. "Neighborhood Interaction in a Homogeneous Community." *American Sociological Review* 15:357-366.

Duneier, Mitchell. 1999. *Sidewalk*. New York: Farrar, Straus and Giroux.

Duque, Juan Carlos, Raúl Ramos, and Jordi Suriñach. 2007. "Supervised Regionalization Methods: A Survey." *International Regional Science Review* 30:195-220.

Festinger, Leon, Stanley Schachter, and Kurt Back. 1950. *Social Pressures in Informal Groups*. Stanford, CA: Stanford University Press.

Fortunato, Santo. 2009. "Community Detection in Graphs." *Physics Reports* In press.

Gans, Herbert J. 1962. *The Urban Villagers*. New York: Free Press.

Grannis, Rick. 2009. *From the Ground Up: Translating Geography into Community through Neighbor Networks*. Princeton: Princeton.

Halperin, Rhoda H. 1998. *Practicing Community: Class Culture and Power in an Urban Neighborhood*. Austin, TX: University of Texas.

Harris, David R. 2001. "Why Are Whites and Blacks Averse to Black Neighbors?" *Social Science Research* 30:100-116.

Hipp, John R. 2007. "Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point." *American Sociological Review* 72:659-680.

Hipp, John R. and Andrew J. Perrin. 2006. "Nested Loyalties: Local Networks' Effects on Neighborhood and Community Cohesion." *Urban Studies* 43:2503-2523.

—. 2009. "The Simultaneous Effect of Social Distance and Physical Distance on the Formation of Neighborhood Ties." *City & Community* 8:5-25.

Hunter, Albert. 1974. *Symbolic Communities*. Chicago: University of Chicago.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities*. New York: Random House.

MapQuest. 2005. "Maps homepage [Web Page]."

Mayhew, Bruce H., J. Miller McPherson, Thomas Rotolo, and Lynn Smith-Lovin. 1995. "Sex and Race Homogeneity in Naturally Occurring Groups." *Social Forces* 74:15-52.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415-444.

Moody, James. 2001. "Peer Influence Groups: Identifying Dense Clusters in Large Networks." *Social Networks* 23:261-283.

Morenoff, Jeffrey D. 2003. "Neighborhood Mechanisms and the Spatial Dynamics of Birth Weight." *American Journal of Sociology* 108:976-1017.

Osgood, D. Wayne and Amy L. Anderson. 2004. "Unstructured Socializing and Rates of Delinquency." *Criminology* 42:519-550.

Park, Robert E. and Ernest W. Burgess. 1921. *Introduction to the science of sociology*. Chicago: University of Chicago.
Paxton, Pamela and James Moody. 2003. "Structure and Sentiment: Explaining Emotional Attachment to Group." *Social Psychology Quarterly* 66:34-47.
Ross, Catherine E., John R. Reynolds, and Karlyn J. Geis. 2000. "The Contingent Meaning of Neighborhood Stability for Residents' Psychological Well-Being." *American Sociological Review* 65:581-595.
Sampson, Robert J. and W. Byron Groves. 1989. "Community Structure and Crime: Testing Social-Disorganization Theory." *American Journal of Sociology* 94:774-802.
Sampson, Robert J. and Stephen W. Raudenbush. 1999. "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods." *American Journal of Sociology* 105:603-651.
—. 2004. "Seeing Disorder: Neighborhood Stigma and the Social Construction of "Broken Windows"." *Social Psychology Quarterly* 67:319-342.
Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277:918-924.
Schwirian, Kent P. 1983. "Models of Neighborhood Change." *Annual Review of Sociology* 9:83-102.
Shaw, Clifford and Henry D. McKay. 1942. *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
Silver, Eric and Lisa L. Miller. 2004. "Sources of Informal Social Control in Chicago Neighborhoods." *Criminology* 42:551-584.
Tatlow, James R., John D. Clapp, and Melinda M. Hohman. 2000. "The Relationship between the Geographic Density of Alcohol Outlets and Alcohol-Related Hospital Admissions in San Diego County." *Journal of Community Health* 25:79-88.
Taylor, Ralph B. 1997. "Social Order and Disorder of Street Blocks and Neighborhoods: Ecology, Microecology, and the Systemic Model of Social Disorganization." *Journal of Research in Crime and Delinquency* 34:113-155.
Taylor, Ralph B., Stephen D. Gottfredson, and Sidney Brower. 1984. "Block Crime and Fear: Defensible Space, Local Social Ties, and Territorial Functioning." *Journal of Research in Crime and Delinquency* 21:303-331.
U.S. Census Bureau. 2005. "American FactFinder [Web Page]." *URL http:/factfinder.census.gov/servlet/AGSGeoAddressServlet? lang=en& programYear=50& treeID=420*.
U.S. Postal Service. 2005. "ZIP code lookup [Web Page]. URL " [*http://zip4.usps.con/zip4/welcome.isp*](http://zip4.usps.con/zip4/welcome.isp)
Wooldredge, John. 2002. "Examining the (Ir)relevance of Aggregation Bias for Multilevel Studies of Neighborhoods and Crime with an Example Comparing Census Tracts to Official Neighborhoods in Cincinnati." *Criminology* 40:681-709.

Tables and Figures

Table 1. Summary statistics for size of network neighborhoods extracted in study of three North Carolina counties

	Number of persons in each neighborhood					Percent overlap with block group with largest overlap			
	Mean	Std Dev	Min	Max	Isolates	Mean	Std Dev	Min	Max
Physical distance	38.2	26.4	1	117	3	0.624	0.215	0.182	1
Social ties X 1	37.9	24.8	1	115	2	0.621	0.205	0.206	1
Social ties X 2	34.1	24.3	1	115	4	0.633	0.216	0.207	1
Social ties X 3	26.4	21.9	1	115	9	0.592	0.247	0.174	1
Multiple neighborhoods	52.5	34.1	1	201	1	0.517	0.198	0.175	1
Block group	31.8	24.2	1	109					
Tract	88.4	93.0	1	291					

Table 2. Odds ratio of tie with someone in same "neighborhood" for three North Carolina counties

	Distance not in model	Distance in model	
	Odds ratio	Odds ratio	Odds ratio for logged distance
Physical distance	4.567	1.269	0.573
Social ties X 1	4.886	1.427	0.584
Social ties X 2	5.302	1.618	0.595
Social ties X 3	8.767	3.776	0.680
Multiple neighborhoods	4.514	0.575	0.551
Block group	4.818	1.837	0.598
Tract	3.850	2.190	0.639

Table 3. Comparing the intra-class correlation of various measures using various constructions of network neighborhoods in all three counties, using ties between adolescents

Parents	Spatial only	Social ties X 1	Social ties X 2	Social ties X 3	Block group	Tract
Crime	0.115	0.125	0.096	0.072	0.120	0.103
Physical disorder	0.051	0.075	0.061	0.038	0.056	0.032
Social disorder	0.103	0.117	0.090	0.076	0.111	0.060
Cohesion	0.092	0.106	0.077	0.065	0.117	0.081
Neighborhood satisfaction	0.093	0.112	0.103	0.095	0.135	0.106
Collective efficacy	0.040	0.044	0.042	0.070	0.064	0.054
Adolescents						
Crime/disorder	0.070	0.075	0.075	0.088	0.080	0.060
Number of neighborhoods	121	123	132	163	132	43

Note: Number of individuals in "Parent" models is 1,133, number of individuals in "Adolescent" models is 3,906.

Figure 1. Hypothetical social network ties if neighborhood effects are present

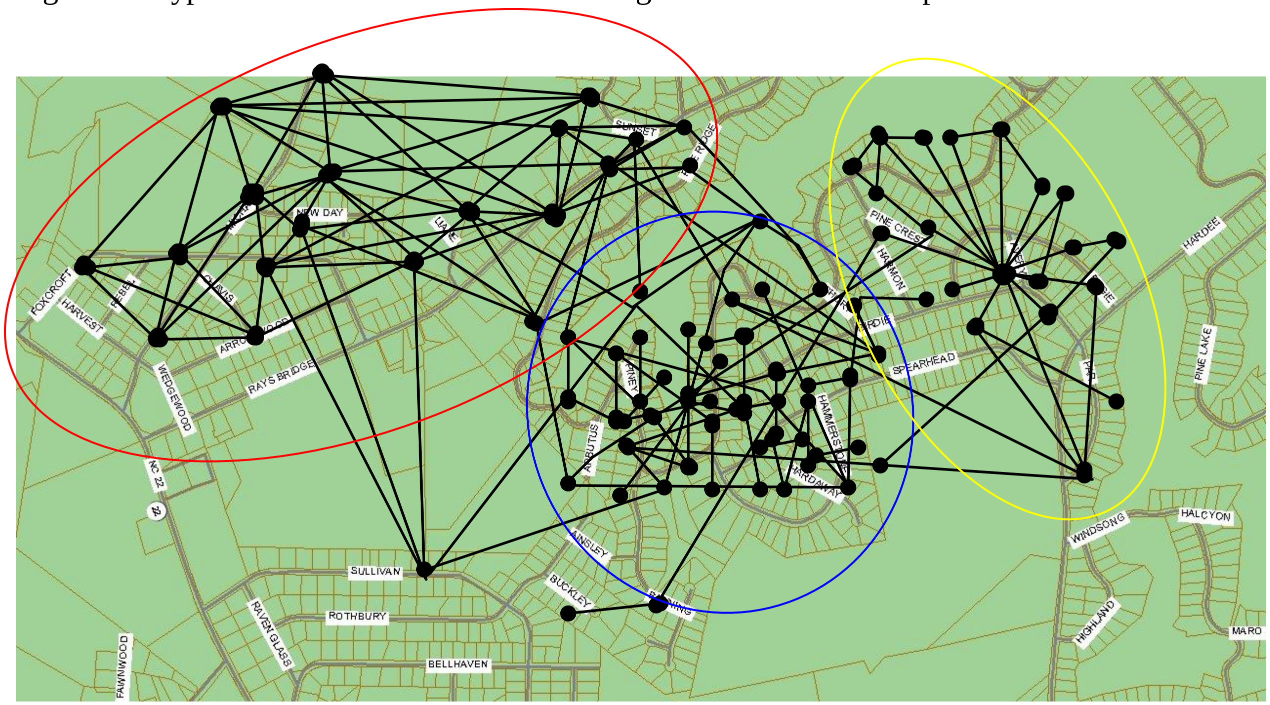
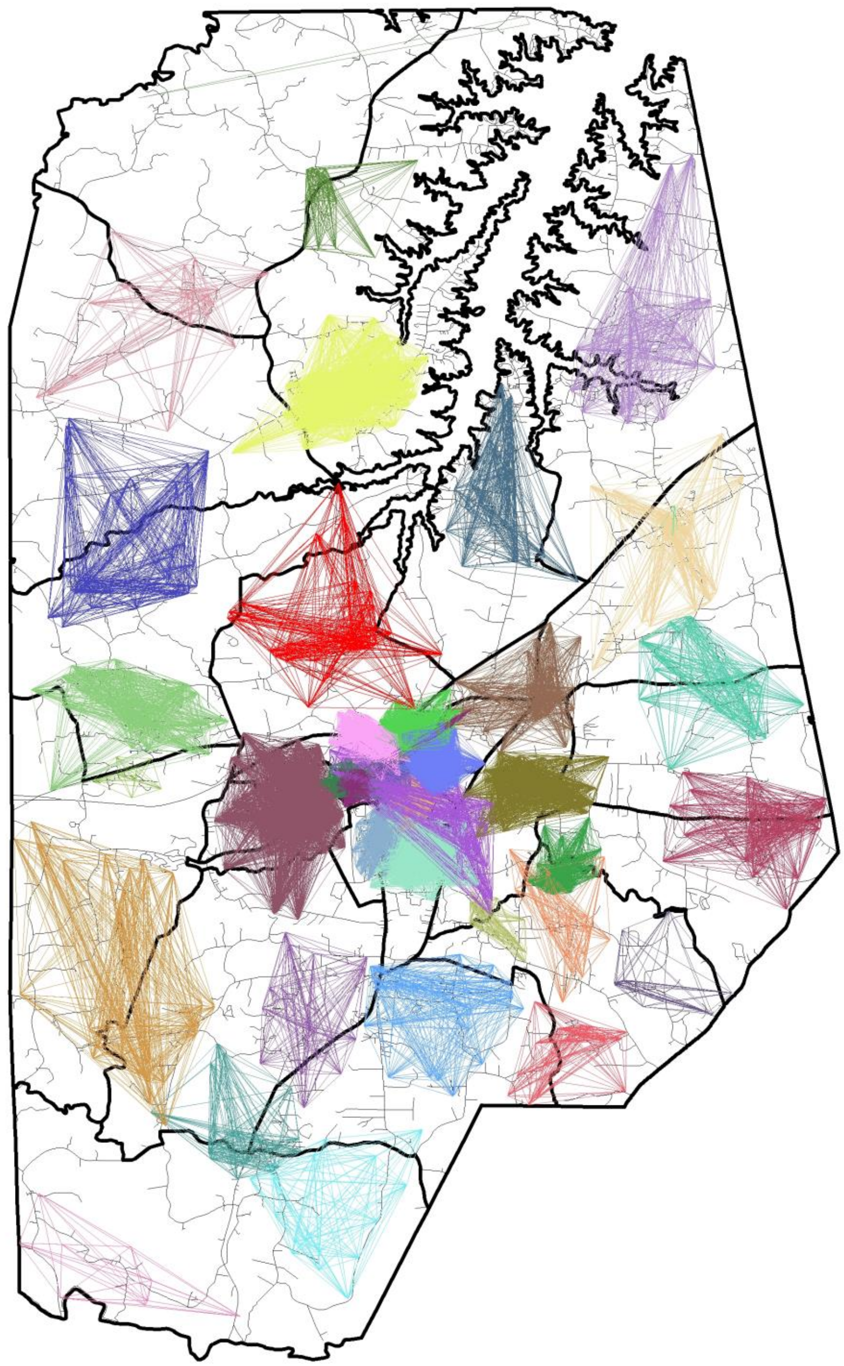
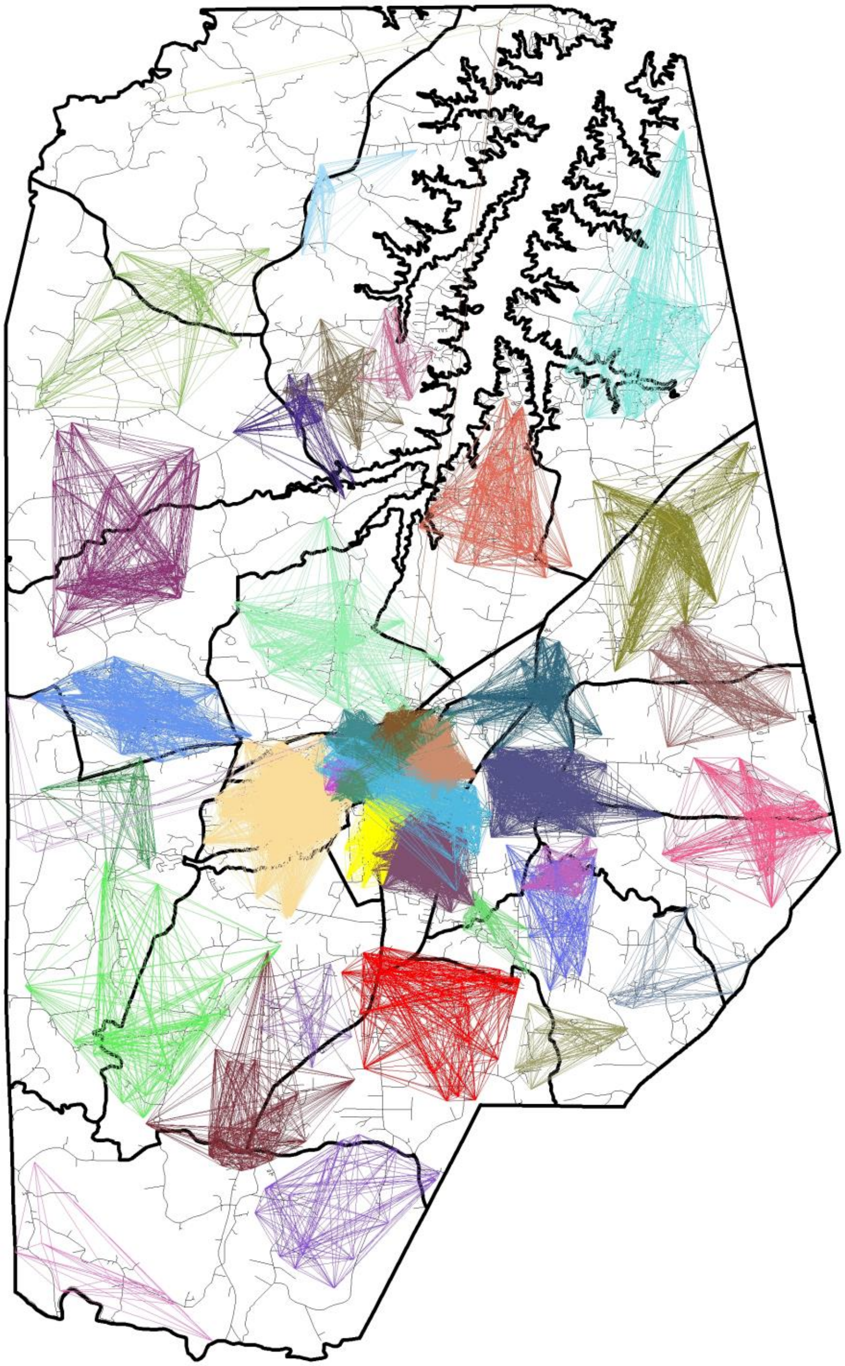


Figure 2. Plotting adolescents in Vance County by network neighborhood (social ties weighted log 1)



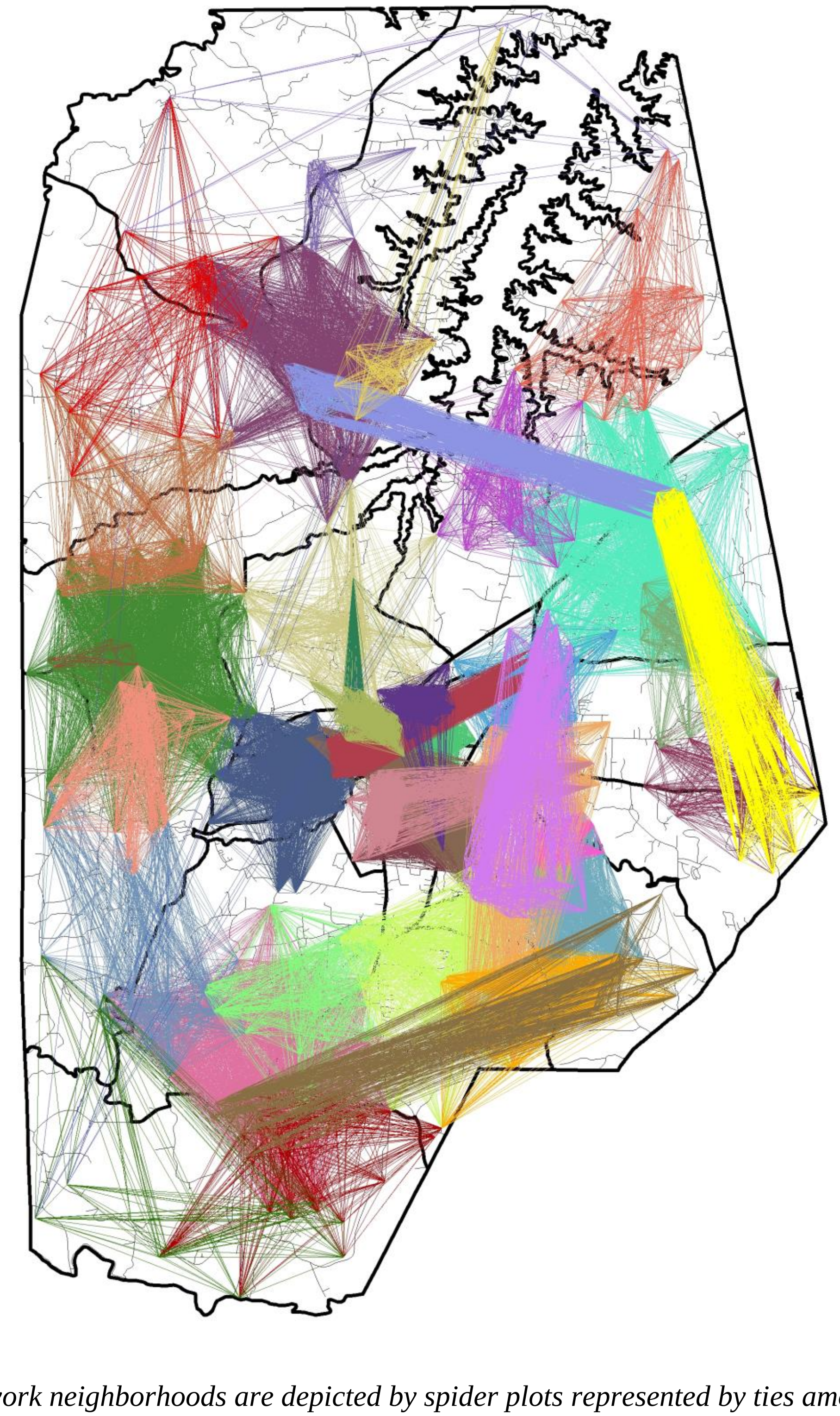
Note: Network neighborhoods are depicted by spider plots represented by ties among all neighborhood members. Block group boundaries are the black lines, and the streets are in gray.

Figure 3. Plotting adolescents in Vance County by network neighborhood, (social ties weighted log 2)



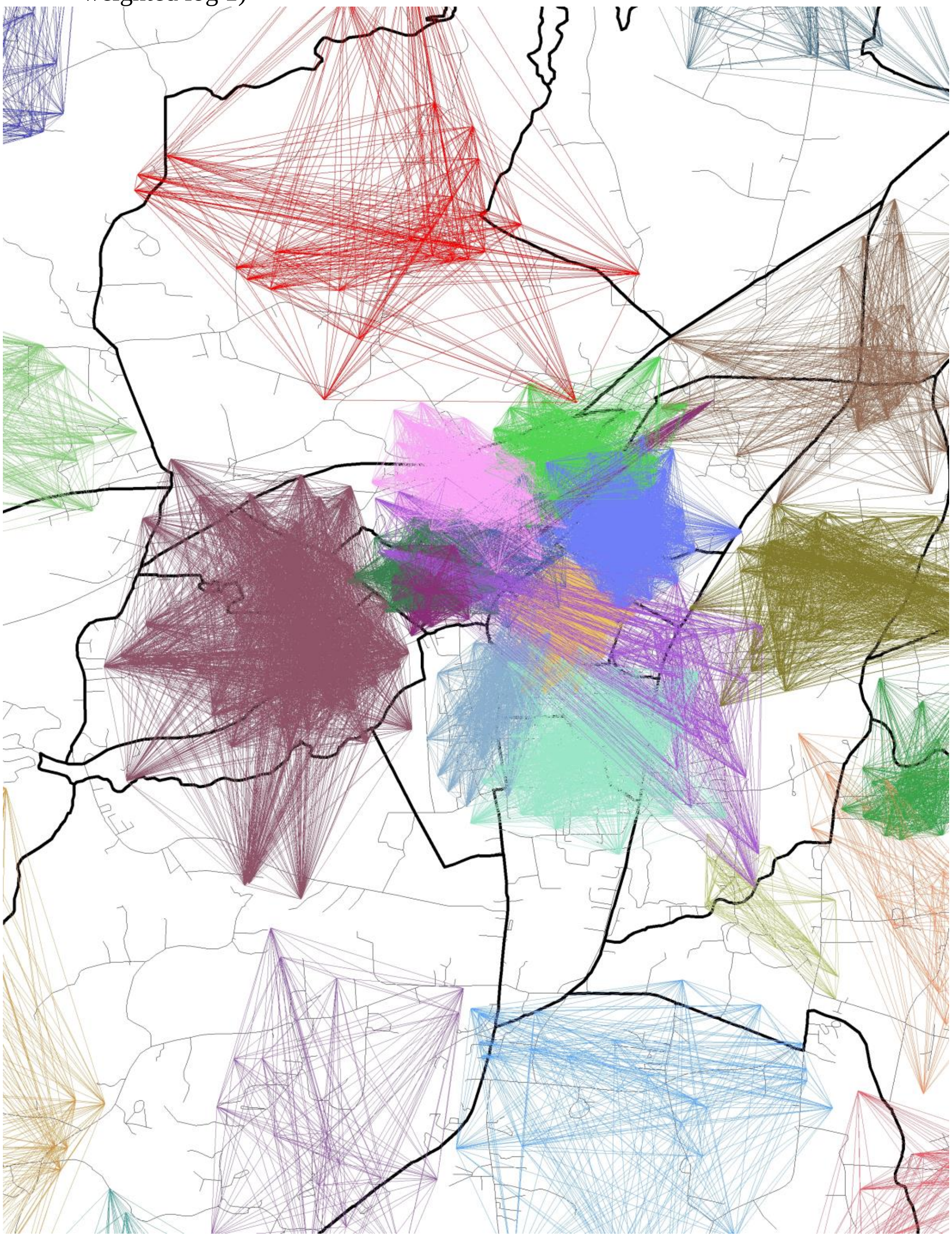
Note: Network neighborhoods are depicted by spider plots represented by ties among all neighborhood members. Block group boundaries are the black lines, and the streets are in gray.

Figure 4. Plotting adolescents in Vance County by network neighborhood, allowing membership in more than one neighborhood



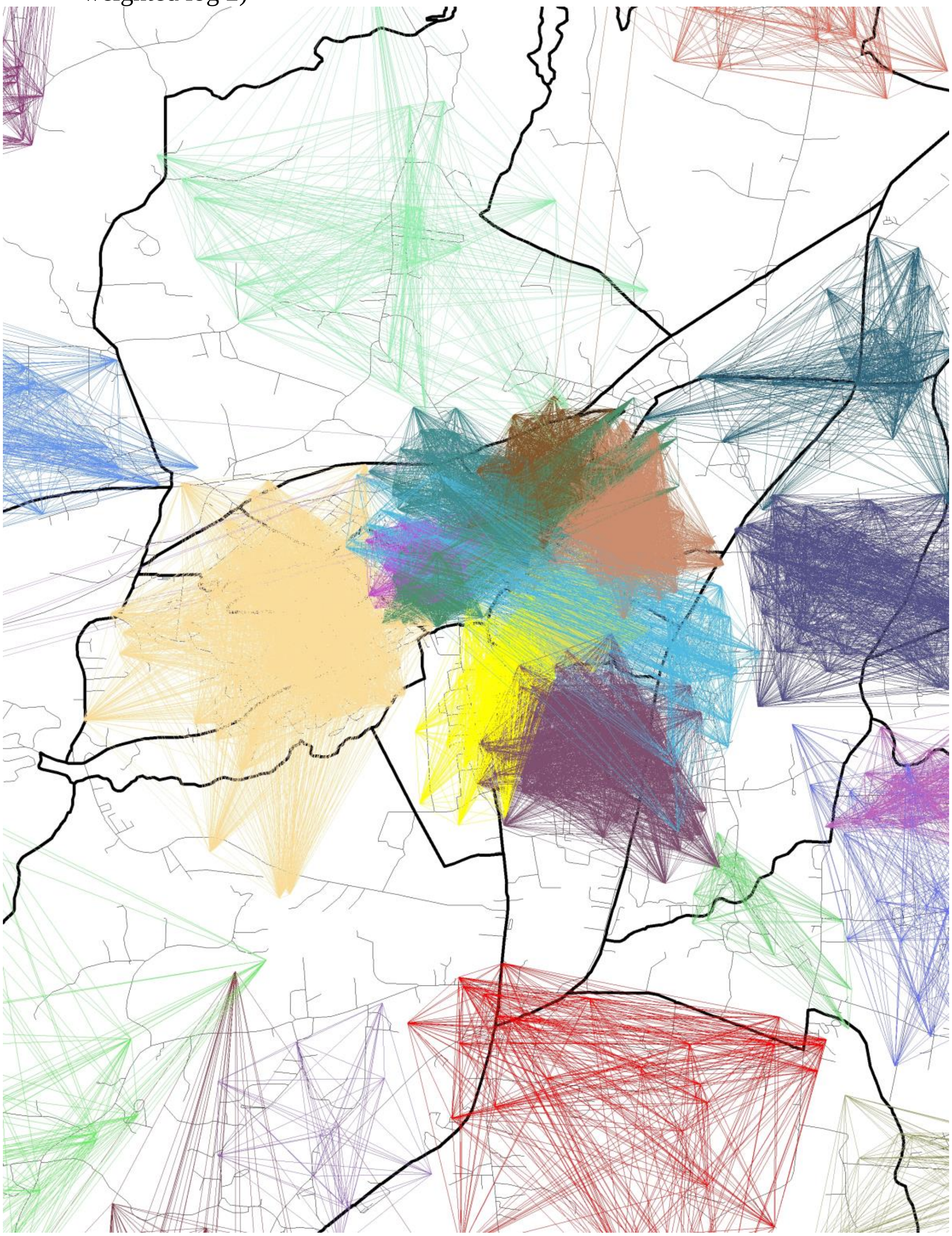
Note: Network neighborhoods are depicted by spider plots represented by ties among all neighborhood members. Block group boundaries are the black lines, and the streets are in gray.

Figure 5. Plotting adolescents in downtown Vance County by network neighborhood (social ties weighted log 1)



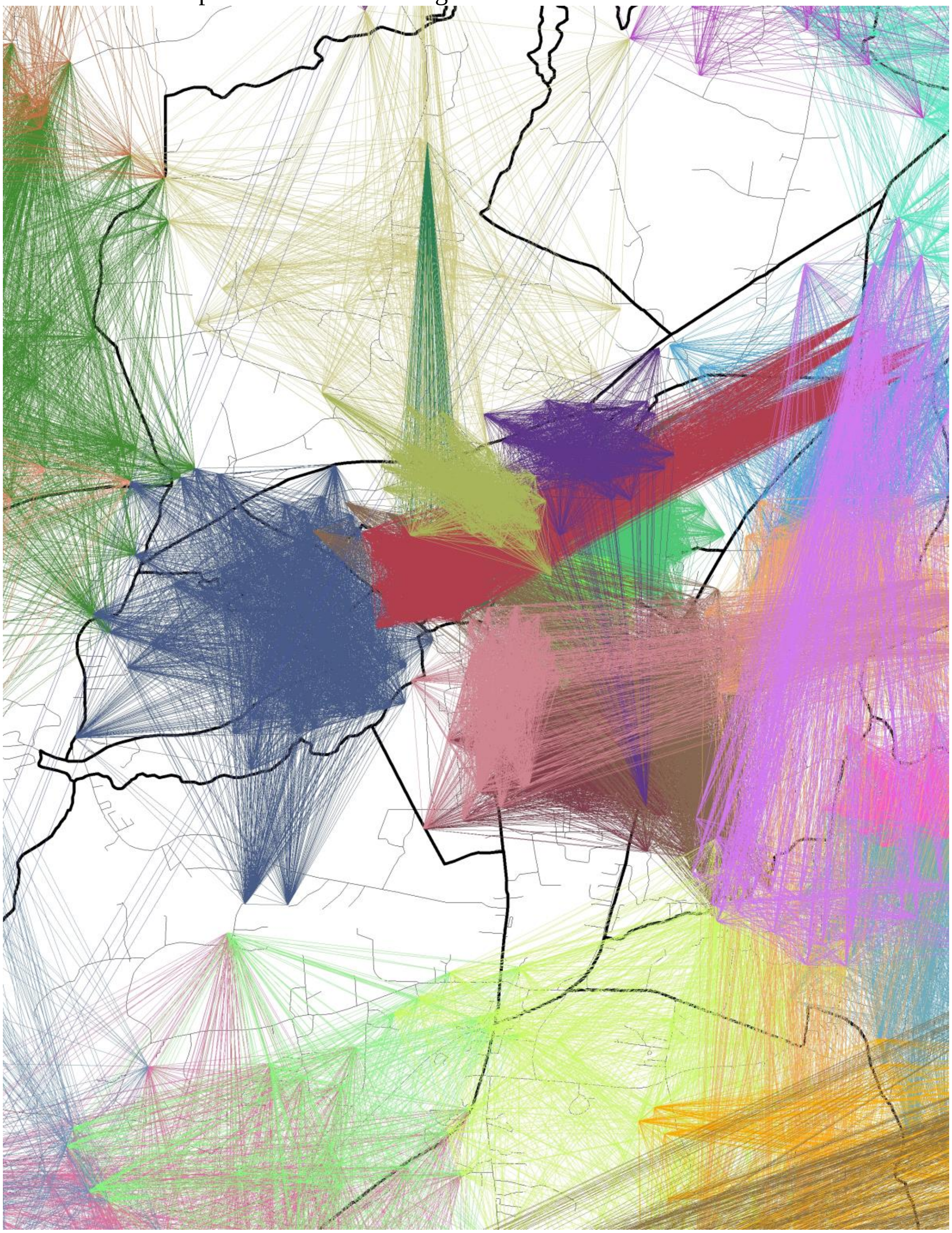
Note: Network neighborhoods are depicted by spider plots represented by ties among all neighborhood members. Block group boundaries are the black lines, and the streets are in gray.

Figure 6. Plotting adolescents in downtown Vance County by network neighborhood, (social ties weighted log 2)



Note: Network neighborhoods are depicted by spider plots represented by ties among all neighborhood members. Block group boundaries are the black lines, and the streets are in gray.

Figure 7. Plotting adolescents in downtown Vance County by network neighborhood, allowing membership in more than one neighborhood



Note: Network neighborhoods are depicted by spider plots represented by ties among all neighborhood members. Block group boundaries are the black lines, and the streets are in gray.