



Gender clustering in friendship networks: some sociological implications

Deirdre M. Kirke

Abstract

This paper uses a social network approach to examine gender clustering in a complete network of teenagers and their friends. It demonstrates the advantages of using increasingly sophisticated social network techniques, including clustering coefficients and their visualization, and social selection models within the ERGM framework, to visualize and explain the process of clustering which takes place in teenagers' networks. The paper supports previous findings of gender homophily among teenagers in small cliques of friends, provides evidence of clustering among larger groups of friends that differs by gender and evidence that the process of clustering also differs by gender. Males make more friends and form larger clusters than females. Differences in clustering are due to differences in selection (males make more friends), triadic closure (more likely for females) and endogenous effects (impacting more on males). These findings have sociological implications for single-gender and cross-gender influences on teenagers' behaviour, and for the presumed importance of agency (selection) over structure (endogenous effects) on friendship formation.

Key words: friendship clustering, gender homophily, friendship networks

Introduction

Friendship studies have consistently confirmed a tendency towards gender homophily in friendship ties and that this is so whether the subjects are children (Eder and Hallinan, 1978; Berenson, Apostoleris and Parnass, 1997), adolescents (Ennett and Bauman, 1996; Pearson et al., 2006) or adults (McPherson, Smith-Lovin and Cook, 2001; Fischer and Oliker, 1983; Campbell and Lee, 1992). The question of whether males and females cluster their friendship ties differently has not been addressed in previous research. It could not have been answered by research which collected data from individuals on the gender of their friends or on dyadic friendship data only. Nor could it have been answered by data on small cliques, usually of about three to five individuals (Ennett and Bauman, 1996; Pearson et al., 2006), because that clustering is done in larger networks, behind the scenes through indirect ties. Individuals themselves would be unaware of the patterning of their friendship ties beyond their egocentric network and the clustering by gender that ensues from that patterning. The question of gender clustering could not have been adequately examined either using conventional research methods, such as surveys or ethnographic methods, and the related data analytical techniques. The question of gender clustering can, however, be examined using a social network approach. It is examined in this paper using a social network approach applied to complete network data and using recently

available social network techniques, including clustering coefficients and their visualization, and social selection models within the ERGM framework, to visualize and explain the process of clustering which takes place in teenagers' networks.

The social network approach is a distinctive way of examining the social world: the theoretical approach is relational, the data are relational, the statistical tests focus on relational properties of networks (Freeman, 2004; Wasserman and Faust, 1994) and mathematical models explore the processes at work in networks of relational ties (Robins, Elliott and Pattison, 2001). The theoretical belief underlying social network research is that the structure, or pattern, formed by the relational ties between individuals in social networks is central to our understanding of the social world (Wellman, 1988; Collins, 1988). The theoretical perspective, at its simplest, is that social structures affect individual actions and that the individuals, in turn, form and change those structures when they form and change relationships. The structures provide access to various resources, such as knowledge, influence, social support and social capital, for their members. Social network analysts believe, therefore, that individual action is best understood in the social context in which it is embedded and that social context can be found in the social structures of relationships formed by the individuals concerned. Those social structures can be described as social networks, which consist of finite sets of actors and the relationships between those actors (Wasserman and Faust, 1994, 20).

Combined with this theoretical perspective is a methodology which enables social network analysts to directly examine the social structures involved, examine how relationships are patterned and examine individuals' actions or social processes in the context of those structures. To facilitate the examination of structures of various size and type, social network data can be examined at many levels, including egocentric networks, dyadic ties, triadic ties, larger partial networks or a complete network. An egocentric network includes an individual's friendship ties and all of the friendship ties between those friends (Knoke and Kuklinski, 1982, 16). Dyadic ties are between two individuals. In the research reported in this paper they are usually between a teenager and a friend. Triadic ties are between three individuals and they are similar to triangles in the ERGM data discussed later in this paper. Partial networks can vary in size. In the discussions of the Kirke data partial networks were the friendship groups identified by the weak component procedure. Weak components were identified in the Kirke data using the GRADAP (Sprenger and Stokman, 1989) programme. A weak component is the maximum unique subset of points which are connected, directly or indirectly, to each other by lines (Harary, Norman, and Cartwright, 1965, 405) (see Scott, 1991, 28-31 for other definitions). Complete networks are sometimes described as whole networks. A complete network is covered when data is collected on all relationships, of a certain kind, existing between all actors within a particular population (Knoke and Kuklinski, 1982, 17). In the Kirke data the actors were the teenagers and the relationships were friendships which existed between the teenagers in the community studied. Although differing greatly from conventional research methods in their relational focus, social network analysts can use any of the conventional data gathering techniques, such as surveys, ethnographic methods and various documents, to collect relational data. Which one is appropriate is decided by the researchers depending on their research problem and the data needed to resolve it.

When I first analyzed my social network data only the programme GRADAP (Sprenger and Stokman, 1989) was available and suitable for the analyses I needed. Many other software programmes have become available in recent years for the analysis of social network data. The most widely used currently are UCINET (Borgatti, Everett and Freeman, 2006) and PAJEK (Batagelj and Mrvar, 2003) (see www.insna.org for list of network software available). I use PAJEK in this paper to provide calculations and visualizations of clustering that were not possible at the time that I collected the data. In recent years researchers have been developing mathematical models from the Exponential Random Graph Modelling framework to examine processes at work in networks in cross-sectional data (Robins, Elliott and Pattison, 2001) and in longitudinal data (Snijders, Steglich, Schweinberger and Huisman, 2005). Since my data are cross-sectional I use the ERGM framework in this paper to model the processes at work in the structuring of the complete network of friendship ties found in this community years earlier (Kirke, 1990). Thus, recent software programmes and mathematical modelling techniques are used in this paper to provide new insights into the question of gender clustering.

The theoretical question being addressed in the research determines the level of social network data needed and the analytical techniques and mathematical modelling tools needed. The theoretical question being

addressed in this paper is whether clustering occurs in teenagers' friendship networks and whether gender or age affects that process. This question requires social network data larger than cliques, ideally complete network data. Such a data set was created by the author in 1987 (Kirke, 1990) to examine the role of friendship networks in substance use. Software programmes were not available at that time to examine questions of clustering nor were the mathematical modelling techniques available to examine the processes at work in the clustering of teenagers' friendship ties. As stated above, they are available now and they are used in this paper to examine whether there is evidence of clustering in friendship networks of teenagers (PAJEK, Batagelj and Mrvar, 2003), and whether the individual attributes of gender and age explain the social selection processes at work in such clustering (Robins, Elliott and Pattison, 2001).

Research Method

The dataset on which this paper is based was collected in 1987 by the author in a Dublin working-class community of approximately 2,500 people. A census of every household in the District Electoral Division (DED) was carried out to locate all of the teenagers of 14-18 years living there. A total of 298 teenagers were located and 267 (89.6%) were successfully interviewed in their own homes.

The teenagers were asked in the survey to name all of their friends and the gender and age of all of them. This friendship data produced the network data. Since the friendship data were collected from a total population of teenagers, friendship data were produced on a complete network of friendship ties. Friends who were named, but were outside the population, were not all interviewed. Some (106) were interviewed using a snowball sampling technique and the others were not interviewed. Whether interviewed or not, data had been collected on the friendship and the gender and age of all of the named friends. The complete network dataset was extended to include all of the named friends and this extended dataset was used to examine gender clustering (Kirke, 2006b). While only 267 teenagers were interviewed in this community, they mentioned 1188 friends. Only 278 of these friends were within the complete network. Thus, these 278 dyads form the complete network and the 1188 dyads create the extended dataset. The extended dataset is used in this paper to examine the nature of the gender clustering and data from the complete network and case studies of peer groups are examined when discussing the sociological implications of gender clustering.

Data used in this paper derive from applying social network methods to survey data to produce multilevel data on individual teenagers, in dyads and in peer groups in a complete network using GRADAP (Sprenger and Stokman, 1989). It also uses PAJEK (Batagelj and Mrvar, 2003), to calculate clustering coefficients to examine the extent of gender clustering and to visualize the gender clustering in the extended dataset (Kirke, 2006b) and uses exponential random graph modelling (ERGM) to identify and explain the process of gender clustering among teenagers in this community (Kirke, Pattison and Robins, 2007). The paper draws on findings presented in a variety of publications to explore the sociological implications of the level of clustering found.

Findings

Gender homophily

A first step in examining the tendency for gender homophily among these teenagers and their friends was to examine gender homophily within the dyads. The findings within the complete network and the extended data set are similar: males predominantly select other males as friends and females predominantly select other females as friends, confirming the results of previous research (Ennett and Bauman, 1996; Pearson et al., 2006). The findings in Table 1 confirm that nearly 90% of males select other males as their friends in the complete network (89.4%) and in the extended dataset (87.4%) and that a similar percentage of females select other females as friends in the complete network (88.9%) and slightly fewer in the extended dataset (79.6%). In both datasets the association was significant ($p < 0.001$ in complete network; $p < 0.0001$ in the extended dataset).

Table 1**Gender homophily in dyads in complete network and extended dataset dyads**

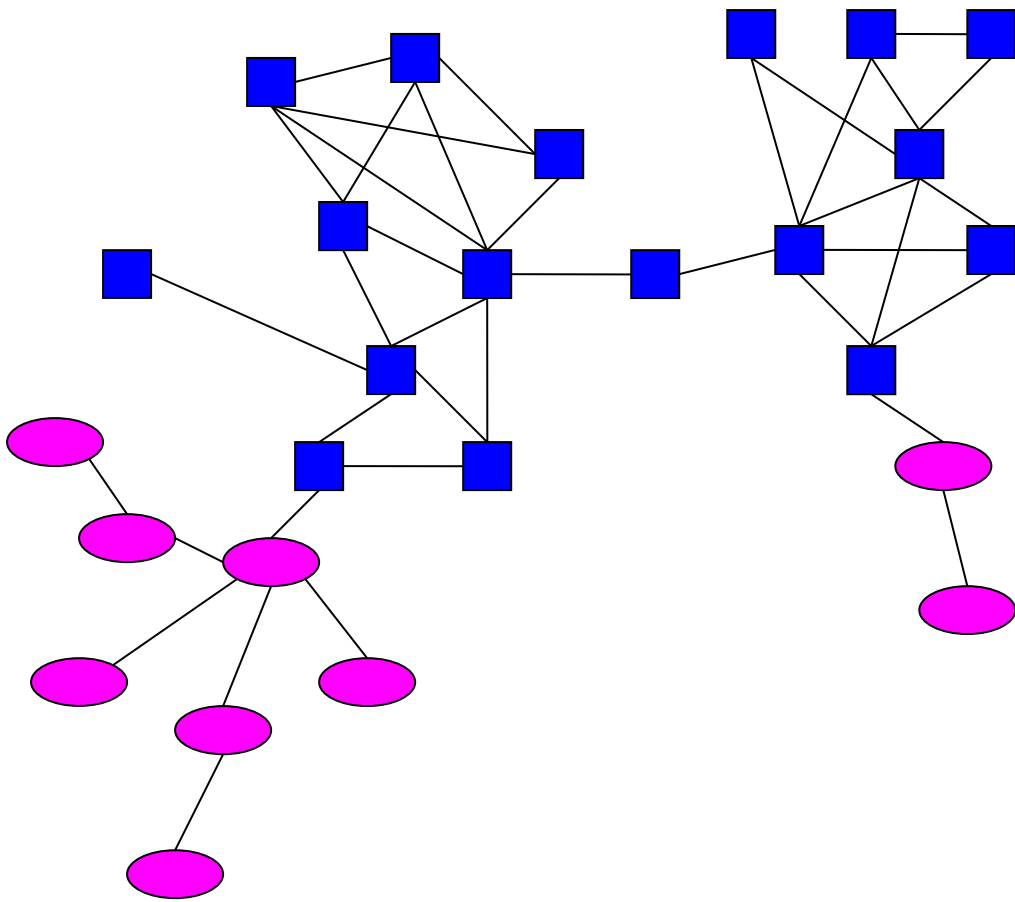
Gender of friend	Gender of teenager In complete network		Gender of teenager in extended dataset	
	Male %	Female %	Male %	Female %
Male	89.4	11.1	87.4	20.4
Female	10.6	88.9	12.6	79.6
Total	100.0	100.0	100.0	100.0
n	170	108	628	560

A second step in exploring the question of gender clustering is to examine how this tendency for homophily in the teenagers' dyadic ties transforms into friendship networks or groups. This is done by using the social network programme GRADAP (Sprenger and Stokman, 1989) to identify the 'weak components' and the procedure 'adjacency' to identify who was connected to whom within each weak component in the complete network. Both procedures occur simultaneously and result in delineating the teenagers' friendship groups in their community (Kirke, 1996, 340). Results confirmed that there were 35 disparate friendship groups within the complete network with a further 98 teenagers having none of their named friends in the complete network. When the attribute gender is added to the friendship groups we find gender homophily in 28 of the groups; 12 all-male groups, ranging in size from ten to two, and 16 all-female groups, ranging in size from 7 to 2. There was a gender mix in the three largest groups of 26 teenagers (17 male: 9 female), 22 teenagers (11m:11f) and of 21 teenagers (9m: 12f), in one group of 3 teenagers (1m: 2f) and in three groups of just 2 teenagers (Kirke, 2006a, 137). While these findings confirm the results of previous research that there is a tendency for gender homophily in teenagers' friendship cliques (Ennett and Bauman, 1996; Pearson et al., 2006), results from the larger networks do not support them. These researchers (Ennett and Bauman, 1996; Pearson et al., 2006) used the definition of clique used by the NEGOPY (Richards, 1989) programme which analyzed their data. Clique members were, according to NEGOPY, adolescents in a group of at least three members, who had most of their links with other members of the same group and were all connected by a path entirely within the group.

Gender clustering

Visualizing the larger networks in the Kirke data will enable us to observe whether there is gender clustering among the males and females in the networks. This is done below for the three largest networks. Figures 1, 2 and 3 add gender to the individuals forming the largest networks of 26, 22 and 21 teenagers respectively. The reader will note that, in Figure 1, the males form two rather dense clusters with the female teenagers forming tree-like structures at the lower end of the figure.

Figure 1



In Figure 2 the two clusters are mixed gender, and there are two female, and one male, tree-like structures at the top of the figure. In the third figure (Figure 3) females form one large dense cluster on the right of the figure and males form a smaller dense cluster on the left of the figure with males and females separately forming tree-like or single ties to other more dense parts of the figure. While it is not possible to reach any firm conclusion about similarities or differences in male and female gender clustering in this community from these findings, it is possible to confirm that some clustering is occurring and it does arouse the curiosity of the researcher to try to find other social network tools to aid in examining the hunch that there is clustering and that different processes may be at work between the males and females.

When preparing a paper for the international network conference in Vancouver, I extended the dataset so that I could examine gender clustering that might involve friends beyond the complete network (Kirke, 2006b). During my participation in a PAJEK workshop prior to the conference, I established that the PAJEK programme could calculate the clustering coefficients by gender for my datasets and furthermore could visualize gender in resulting clusters. PAJEK is a software programme which can analyse very large networks and can produce excellent visualizations of same. With the help of Jurgen Pfeffer we established the nature of that clustering. The results are given below in Figures 4 and 5.

They are given separately for the teenagers in the complete network and all of their named friends who were interviewed, either within or outside the complete network (Figure 4), and for the teenagers in the complete network and all of the friends they named, whether they were interviewed or not (extended dataset) (Figure 5). The colours distinguish the males (blue) from the females (pink), the size of the disks denote the size of the clusters and lines indicate links between the clusters.

Figure 4

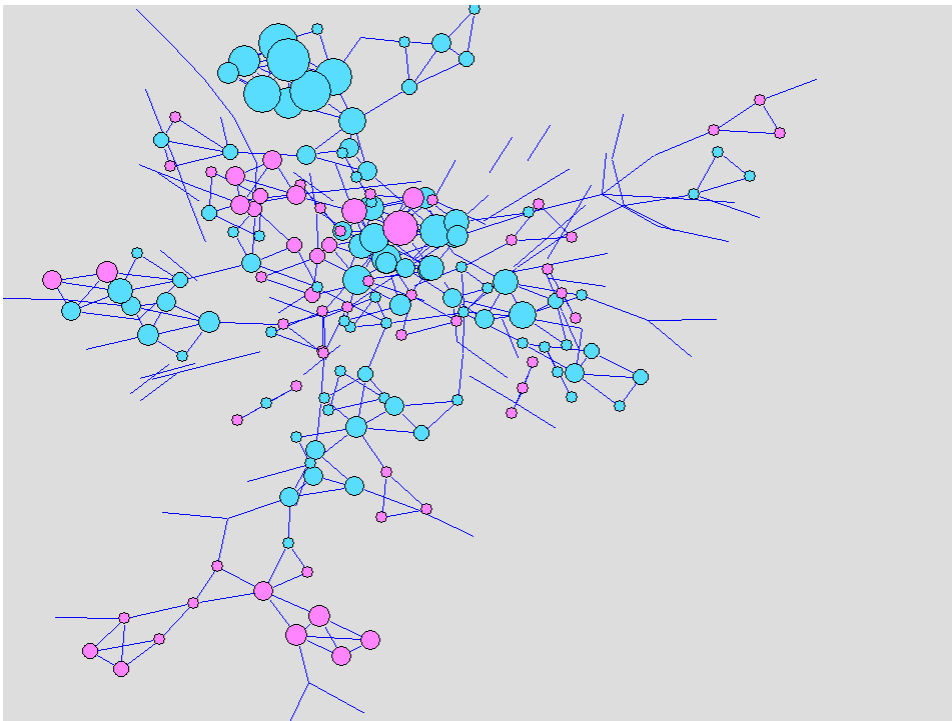
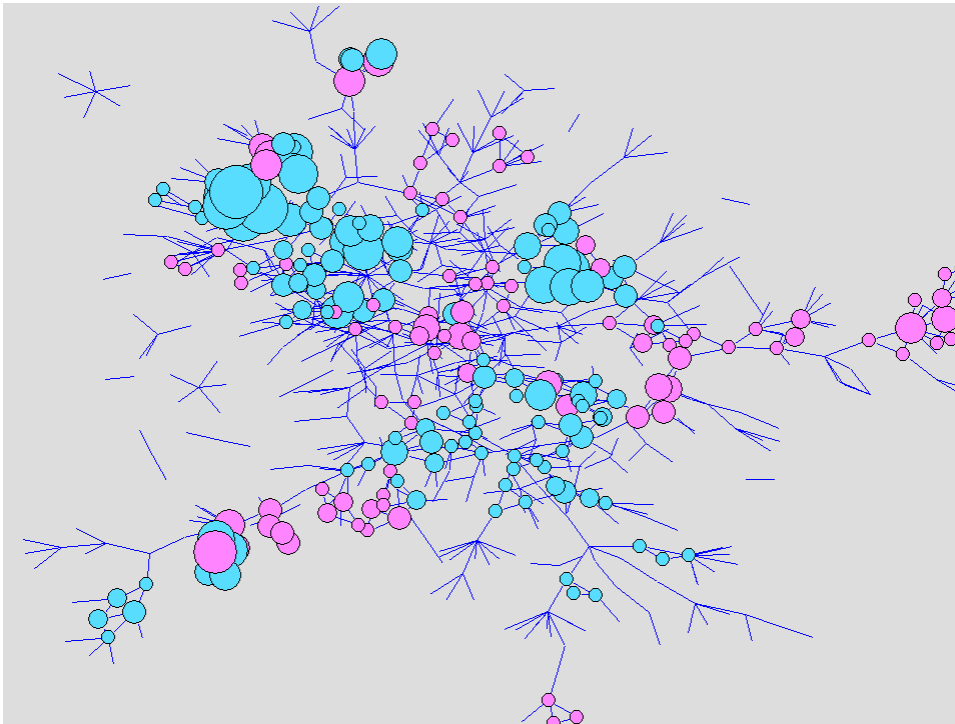


Figure 5

The patterns of clustering are the same in both figures. The size of the clusters is quite different for males and females. The larger clusters are predominantly male clusters. Throughout the figure there are direct lines linking clusters of males with males and females with females. These findings suggest that at a later time, or if the friendship data were more complete, the clusters would become larger as those ties link clusters together but this would not change the relative size of male and female

clusters. Male clusters would continue to be the larger clusters. Figure 5 confirms those results. Indeed it demonstrates that, when all named friends are included, there are larger clusters of both males and females but the largest clusters are still the male clusters. These results suggest that males in this community made lots of friends and that they extended those friendship groups by forming friendships with those to whom they were indirectly tied, thereby enlarging their friendship groups. Females, on the other hand, made fewer friends and, while they enlarged their friendship groups by forming friendships with those to whom they were indirectly tied, they appeared to favour smaller rather than larger clusters of friends.

The process of clustering

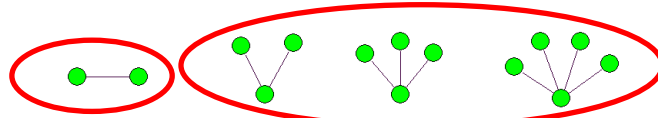
Following my presentation of these findings, Philippa Pattison suggested that it would be possible to model the process of clustering that was at work in this community which resulted in such apparent clustering by gender and we did this for the international network conference in Corfu in 2007 (Kirke, Pattison and Robins, 2007). We used social selection models, within the exponential random graph modelling (ERGM) framework, to model the clustering within the network of friendship ties within the community (see *Social Networks*, 2007, 29, 2 for special section on 'Advances in Exponential Random Graph (p^*) Models'). By clustering we mean whether a teenager's friends' friends were likely to become friends with the teenager in question. ERGM models are used to model social network structure and social selection models assume that individuals select others as friends, and that they are selected as friends, on the basis of some attributes (Robins, Elliott and Pattison, 2001).

Figure 6

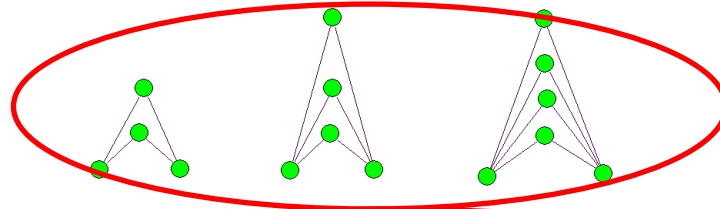
Exponential random graph models for networks

We model the probability of a network in terms of the propensity for certain *families* of network configurations to occur

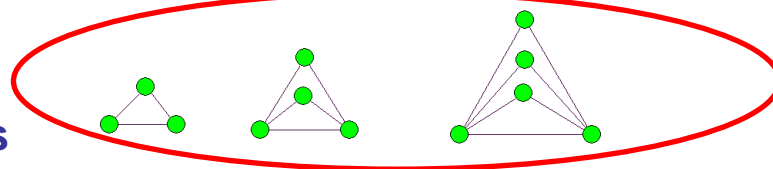
Edges



Stars

(k-stars)






Multiple paths



A good model needs to be estimable from observed data, be a reasonable representation of those data, be theoretically plausible about the type of effects, e.g. individual attributes, which might have produced the network, and be amenable to examining which competing effects might be the best explanation of the data (Robins and Morris, 2007, 169). Only teenagers, who had been interviewed ($n=355$ (18 incomplete data)), were included in this analysis. The two attributes gender and age were used in this analysis to examine their role in structuring the teenagers' networks and the clustering observed within this community. They were used as exogenous actor-level covariates in the social selection model (Robins, Elliott and Pattison, 2001). The models were fitted using PNet (Wang et al., 2006).

On the network side the extent to which different families of network patterns occurred was modelled. These network patterns included edges, stars (k -stars), multiple paths (k -2-paths) and multiple triangles (k -triangles) (Figure 6). The first social selection model examined the extent to which there was differential participation in those network structures by the teenagers in this community in order to model the extent of clustering. Results from Model 1 confirmed that the overall network of friendship ties was sparse but with a high degree of clustering, which was distributed across the network rather than being highly centralized. It also suggested that there was a very weak heterogeneity in the number of friendship ties.

Model 3: adding a gender/triangle interaction

<i>Parameter</i>	<i>estimate</i>	<i>est. s.e.</i>
Edge:	-4.480	0.570
Isolates:	-0.095	0.366
<i>k</i> -Star:	0.118	0.223
<i>k</i> -triangle:	1.923	0.112
<i>k</i> -2-path:	-0.272	0.060
gender homophily: 	3.592	0.434
differential activity: 	-1.522	0.165
age difference: 	-0.266	0.035
all female 2-star: 	-0.435	0.119
differential clustering 	0.269	0.093

Gender and age were added in Models 2 and 3. The results given in Table 2 give the final results of the modelling process (Model 3). The results are read by comparing the estimate with the standard error of the estimate. If the estimate is equal to, or more than, twice the standard error of the estimate, we can be confident about the importance of the result. A positive estimate means that you are seeing more of these in the network than you would expect to see by chance and a negative estimate means that you are seeing fewer of these in the network than you would expect. Reading these results we conclude that there were much fewer edges in this community than might be expected, that is, many fewer dyads of friends who had no other ties. There were more *k*-triangles, that is triadic closure was more likely to happen than might be expected, and much fewer *k*-2-paths, that is, triadic closure had not happened, than might be expected. There were regions high in clustering across the whole community. There was much more gender homophily and, correspondingly, less differential gender activity, observed than might have been expected. Fewer all female 2-star patterns than might have been expected were observed, suggesting that females were more likely to have triadic closure. The final result confirms this further since it confirms that, if even one female were in the triad, there was more likely to be triadic closure. The model also found that there was much less age difference than expected among friends.

Conclusions from the modelling process are that the males had more friendship ties than the females; that friendship ties were more likely to be between teenagers of the same gender and closer in age; that all-female two-path patterns were less likely to be observed and, correspondingly, that triadic closure was stronger for triads in which there was at least one female; and that there were strong endogenous network influences at work which yielded regions high in clustering across this community. Thus, while some of the clustering was explained by the social selection of friends on the basis of their gender or age, some clustering occurred on the basis of endogenous influences. Some of the endogenous influences which would have led to such clustering include the teenagers living next door, on the same road or in the neighbourhood as other children of about their age, attending the same school and being in the same class, and, due to the low income of the families, spending most of their free time in this community.

Thus, the findings on gender clustering visualized by PAJEK in Figures 4 and 5 are confirmed and explained further by these social selection models. The large differences in gender clustering observed in this community in Figures 4 and 5, was due to the tendency for gender homophily in friendship ties, the larger number of ties made by males and endogenous effects impacting more on males than on females. Females, on the other hand, also tended towards gender homophily in their friendship ties, and they were more likely to

have triadic closure but, with their tendency for fewer friends, they were likely to form smaller clusters and, it would appear that the endogenous effects would not have impacted on them as much as on the males.

Sociological implications for these findings

Having established above that there was considerable clustering in this community of teenagers, that males were predominantly in male-only clusters and that females were predominantly in female-only clusters, that male clusters were usually larger and that teenagers were likely to be clustered with others close to their age, this section explores some sociological implications of these results.

An association between the substance use of teenagers and their friends has been established for many years. Using longitudinal data on best friend dyads, Kandel (1973) established that teenagers and their friends were likely to be similar in their drug behaviour and, furthermore, that teenagers' friends were the main source of influence. These results coined the concept 'peer influence'. Research in the 1970s and 1980s consistently confirmed Kandel's results (Akers et al., 1979; Dembo, Schmeidler and Burgos, 1982; Brook, Whiteman and Gordon, 1983). While there was no direct evidence in these studies that particular peers had influenced particular teenagers into using substances, there was evidence in the longitudinal data that teenagers, whose peers had taken the substance at one point in time, were likely to have taken a similar substance at a subsequent time point. On the basis of these findings, the inference drawn was that their peers had influenced the teenagers into substance use.

Later other researchers used a social network approach to establish the level of similarity in the substance use of teenagers in cliques and how this level of similarity could be explained. Many researchers established similarity in one or more of the three substances, cigarettes, alcohol or drugs, among teenagers in cliques and that the similarity was due to both influence and selection (Hunter, Vizelberg and Berenson, 1991; Ennett and Bauman, 1994; Pearson and West, 2003). None of these researchers, however, had data on networks larger than cliques.

When such questions were examined in networks larger than cliques, the results were somewhat different (Kirke, 2004). When three case studies of networks of over twenty teenagers were examined, the results confirmed that some teenagers in each of these networks did become similar to each other in their substance use. Their similarity was due to peer influence for some, to peer selection for others and to neither peer influence nor selection for others (Kirke, 2004). That paper also confirmed that the peer network, and the patterning of friendship ties in it, had contributed to the similarity by bringing together those who were similar with those who were not, thus enabling social influence to happen. Kirke (2004) concluded that a chain reaction explanation was more appropriate than one based on peer influence and or peer selection since the chain reaction explanation included the role played by the patterning of friendship ties in the network.

Later Kirke (2006a) established that gender impacted on all three aspects of the chain reaction process: the selection of peers, the patterning of peer ties and peer influence. Teenagers selected their peers strictly along gender lines, resulting in almost all single gender peer ties and peer groups. The patterning of ties also reflected a gender influence, with different patterning emerging between males and females suggesting that there was more dense clustering among the males (see Figures 1-3 above). Peer influence, on the other hand, operated differently for male and female teenagers. Male teenagers had been influenced predominantly by other males while female teenagers had been influenced predominantly by males and females (Kirke, 2006a).

Having confirmed earlier in this paper that male and female teenagers did indeed cluster their friendship ties differently, the results have implications for this research on teenage substance use and on related sociological issues.

1. Since males and females clustered their friendship ties into separate gender networks, possible influences on their behaviour would be confined to that gender. Thus, males could only influence males and females influence females. When mixed gender influence was needed, as it was by females, it could only occur in the relatively small number of mixed gender ties available. With the level of clustering observed in this network using recent methodological innovations,

influences could flow very easily among males, although only within their clusters, but would reach quite a large number of teenagers because their clusters were inclined to be large. Influences could also flow very easily within the female clusters but they were inclined to be small, so the influences would not reach so many teenagers. Cross-gender influence, which was needed by females for most of their substance use, was likely to be minimal.

2. The results on clustering have implications for research on dating patterns, sexual behaviour and the spread of sexually transmitted diseases among young people because these patterns of behaviour will only become an issue to the extent that teenagers move out of single-gender clusters and into mixed gender friendship/ dating ties. See Bearman, Moody and Stovel (2004) for very different results on networks recording dating patterns.
3. The level of clustering signifies cohesive networks in which, according to previous research, homophily in behaviour is likely to occur (McPherson and Smith-Lovin, 1987; McPherson, Smith-Lovin and Cook, 2001; Homans, 1950; Collins, 1988) due to the selection of like others or the influence of unlike others. Results from the ERGM social selection models, given above on the clustering process, call these findings into question, because, while gender and age were attributes which explained some of the clustering process, some of it was explained by endogenous forces in which social selection had not taken place at all. Results from Kirke (2004) also questions the level of homophily in networks is likely to be exaggerated when researchers do not have access to network data beyond the clique.
4. A related point is that the ERGM social selection models' finding, that selection as such did not always happen in the formation of friendship ties, calls into question the constantly held assumption in friendship studies, and particularly in substance use research, that teenagers select their friends. As was reported above, the formation of some friendship ties happened without selection occurring. Endogenous influences, such as living next door to each other, living on the same road, or being in the same class in school, explained the formation of many of the ties and, indeed, contributed to the level of clustering.
5. These findings call into question the relative importance of agency and structure in establishing friendship ties. While it has been assumed that individuals exercise agency by selecting others as friends, the ERGM modelling results given above confirmed that endogenous influences, which were structural, rather than agency, were at work in establishing many friendships.
6. It is important to establish whether such levels of gender clustering are replicated in other studies of teenagers in complete network data and whether they remain the same as teenagers mature into adults. There may be an age point when the single-gender clustering changes to allow for a more open pattern of mixed gender ties if romantic relationships, partnerships and marriages are to flourish. Future longitudinal research could usefully focus on such issues.

Conclusions

Readers may suggest that the Kirke data, on which most of the findings given in this paper are based, were unique and that, therefore, they provide an esoteric example of friendship that is unlikely to be found elsewhere. That may be so but, when selecting this community for the Dublin study it was not unique, but was similar to numerous other electoral areas in Dublin on a variety of socio-economic variables. How similar it was in clustering levels can not, obviously, be ascertained. What appears to be unique about it is that it was a complete network of teenagers in a community which was covered by the study and not schools of teenagers as is usually the case (Kandel, 1973; Ennett and Bauman, 1996; Pearson and West, 2003; Pearson and Michell, 2000, Fang et al., 2003; Bearman, Moody and Stovel, 2004). Comparative studies based on communities of teenagers would be greatly welcomed.

The methodological innovations described in this paper, particularly the use of a social network approach, using PAJEK to calculate clustering coefficients and visualize the results, and using social selection models within the ERGM framework to establish the processes at work in the clustering would appear to have applications beyond teenagers' friendship ties and teenagers' substance use. They could, for example, be valuable in organizational research, in research on crime and deviance, in research on the diffusion of innovations, search processes and social capital.

References

- Akers, R. L., Krohn, M. D., Lanza-Kaduce, L. and Radosevich, M. (1979) 'Social learning behavior: A specific test of a general theory', *American Sociological Review*, 44: 635-655.
- Batagelj, V. and Mrvar, A. (2003) PAJEK – Program for large network analysis (See www.insna.org software).
- Bearman, P. E., Moody, J. and Stovel, K. (2004) 'Chains of affection: The structure of adolescent romantic and sexual networks', *American Journal of Sociology*, 110, 1, 44-91.
- Berenson, J. F., Apostoleris, N. H. and Parnass, J. (1997) 'Age and sex differences in dyadic and group interaction', *Developmental Psychology*, 33, 3, 538-543.
- Borgatti, S. P., Everett, M. G. and Freeman, L. C. (2006). UCINET 6 Social network analysis software (see www.insna.org software).
- Brook, J. S., Whiteman, M. and Gordon, A. S. (1983) 'Stages of drug-use in adolescence - Personality, peer and family correlates', *Developmental Psychology*, 19: 2: 269-277.
- Campbell, K. E. and Lee, B. A. (1991) 'Name generators in surveys of personal networks', *Social Networks*, 13: 203-221.
- Collins, R. (1988) *Theoretical Sociology*, San Diego: Harcourt Brace Jovanovich, Inc.
- Dembo, R., Schmeidler, J. and Burgos, W. (1982) 'Processes of early drug involvement in three inner-city neighborhood settings', *Deviant Behavior*, 3: 4: 359-383
- Eder, D. and Hallinan, M. T. (1978) 'Sex differences in children's friendships', *American Sociological Review*, 43, 2, 237-250.
- Ennett, S. T. and Bauman, K. E. (1994) 'The contribution of influence and selection to adolescent peer group homogeneity: The case of adolescent cigarette smoking', *Journal of Personality and Social Psychology*, 67: 4: 653-663.
- Ennett, S. T. and Bauman, K. E. (1996) 'Adolescent social networks: School, demographic, and longitudinal considerations', *Journal of Adolescent Research*, 11: 2: 194-215.
- Fang, X., Li, X., Stanton, B. and Dong, Q. (2003) 'Social network positions and smoking experimentation among Chinese adolescents', *American Journal of Health Behavior*, 27: 3: 257-267.
- Fischer, C. S. and Oliner, S. J. (1983) 'A research note on friendship, gender, and the life cycle', *Social Forces*, 62, 124-132.
- Freeman, L. C. (2004) *The development of social network analysis: A study in the sociology of science*. Vancouver, Empirical Press.
- Harary, F., Norman, R. Z. and Cartwright, D. (1965) *Structural models: An introduction to the theory of directed graphs*. New York: Wiley.
- Homans, G. C. (1950) *The Human Group*, New York: Harcourt.
- Hunter, S. Mac D., Vizelberg, I. A. and Berenson, G. S. (1991) 'Identifying mechanisms of adoption of tobacco and alcohol use among youth: The Bogalusa heart study', *Social Networks*, 13: 91-104.
- Kandel, D. B. (1973) 'Adolescent marijuana use: Role of parents and peers', *Science*, 181: 1067-1070.
- Kirke, D. M. (1990) *Teenage Drug Abuse: An Individualistic and Structural Analysis*, Ph.D. Thesis, University College Dublin.
- Kirke, D. M. (1996). 'Collecting peer data and delineating peer networks in a complete network', *Social Networks*, 18: 4: 333-346.
- Kirke, D. M. (2004) 'Chain reactions in adolescents' cigarette, alcohol and drug use: Similarity through peer influence or the patterning of ties in peer networks?', *Social Networks*, 26: 1: 3-28.
- Kirke, D. M. (2006a) *Teenagers and substance use: Social networks and peer influence*. Basingstoke and New York: Palgrave Macmillan.
- Kirke, D. M. (2006b) 'Determinants of gender composition and structure of peer groups in a community'. Paper presented at the International Sunbelt Social Network Conference, Vancouver, April.
- Kirke, D. M., Pattison, P. and Robins, G. (2007) 'Modelling the role of gender in structuring adolescent friendship networks'. Paper presented at the International Sunbelt Social Network Conference, Corfu, May.
- Knoke, D. and Kuklinski, J. H. (1982) 'Network Analysis', *Sage University Paper Series on Quantitative Applications in the Social Sciences*, Beverly Hills and London, Sage Publications: Series No. 07-028.

- McPherson, J. M. and Smith-Lovin, L. (1987) 'Homophily in voluntary organizations: Status distance and the composition of face-to-face groups', *American Sociological Review*, 52: 370-379.
- McPherson, J. M., Smith-Lovin, L. and Cook, J. M. (2001) 'Birds of a feather: Homophily in social networks', *Annual Review of Sociology*, 27: 415-444.
- Pearson, M. and Michell, L. (2000) 'Smoke rings: Social network analysis of friendship groups, smoking, and drug-taking', *Drugs: Education, Prevention and Policy*, 7: 1: 21-37.
- Pearson, M. and West, P. (2003) 'Drifting smoke rings: Social network analysis and Markov processes in a longitudinal study of friendship groups and risk-taking', *Connections*, 25: 2: 59-76.
- Pearson, M., Sweeting, H., West, P., Young, R., Gordon, J. and Turner, K. (2006) 'Adolescent substance use in different social and peer contexts: A social network analysis', *Drugs-Education Prevention and Policy*, 13, 6, 519-536.
- Richards, W. D. (1989) *The NEGOPY network analysis program*. School of Communication, Simon Fraser University, Burnaby, BC, Canada.
- Robins, G., Elliott, P. and Pattison, P. (2001) 'Network models for social selection processes', *Social Networks*, 23, 1-30.
- Robins, G. and Morris, M. (2007) 'Editorial: Advances in exponential random graph (p^*) models', *Social Networks*, 29, 169-172.
- Scott, J. (1991). *Social Network Analysis: A Handbook*, London: Sage Publications Ltd.
- Snijders, T. A. B., Steglich, C. E. G., Schweinberger, M. and Huisman, M. E. (2005) *Manual for SIENA version 2.1*. Groningen, ICS.
- Sprengrer, C. J. A. and Stokman, F. N. (Eds) (1989) *Gradap: Graph Definition and Analysis Package*, Groningen, ProGamma.
- Wang, P., Robins, G. and Pattison, P. (2006) *Pnet: A program for the simulation and estimation of exponential random graph models*. University of Melbourne.
- Wasserman, S. and Faust, K. (1994) *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
- Wellman, B. (1988). 'Structural analysis: From method and metaphor to theory and substance', in Wellman, B. and Berkowitz, S. D. (Eds), *Social Structures: A Network Approach*. Cambridge, New York: Cambridge University Press, 19-61
- Wellman, B. and Berkowitz, S. D. (Eds) (1988). *Social Structures: a Network Approach*, Cambridge, New York: Cambridge University Press.

Acknowledgements

I would like to acknowledge the work done by Jurgen Pfeffer in creating the visualizations given in Figures 4 and 5 and Philippa Pattison and Garry Robins in modelling the process at work in clustering in this community.