Social networks in CAVES

methodological overview

Joanna Stefanska

Environmental Research Unit, Faculty of Psychology, Warsaw University

jstefanska@psych.uw.edu.pl

Keywords: social networks, model of social structure, CAVES, sociogram, matrix, sampling

Warszawa-Wrocław-Brussels-Tilburg-Manchester 2005/2006

Introduction

The following report from research on social networks is based on understanding that CAVES project is focused on a general research question: "How do farmers make their decisions about land use in response to external shocks?" The fact, that there is need for a social part relates to the basic assumption that social impact (conceptualised as a social network) is an important factor while deciding about land use. Farming is in fact a kind of business activity and as such is vulnerable to the influence of social environment. Several studies provided arguments that social networks facilitate eg. self-employment (Allen, 2000), business contacts and frauds (Baker & Faulkner, 2004). This effect is explained in terms of social networks providing social support, enhancing information flow, providing access to potential clients, capital and other resources (Allen, 2000).

It is also noteworthy that CAVES becomes one of the first attempts to combine social network data with other types of models (KnETs and GIS data). The earlier attempt by Faust, Entwisle, Rindfuss, Walsh and Sawangdee (1999) was focused on combining social network data with GIS spatial information and provided unique insight into spatial constraints of social contacts.

However, it must not be forgotten that social impact is most probably only one of many factors influencing decision making. Therefore, the social part of CAVES will consist of several approaches combined into a model of social factors in land use change.

- Phase 1: creating a structural model of a social network based on field data.
- Phase 2: re-constructing decision-making algorithms for agents (based on field data, KnETs)
- Phase 3: dynamic simulation of changes in decision making as a function of social factors

Note: This text deals only with the issue of social networks (phase 1).

Social network model

Social network is a model of social structure recording relations between actors. As such, it maybe depicted as a matrix (agents by agents with numbers in cells that represent existence and/or strength of relations) or as a graph consisting of nodes which represent agents and links/edges that represent relations (Scott, 2000). Possible nodes are: individuals, households, villages, municipalities, organizations. Possible links may represent: information exchange, mutual help, co-operation, example giving, convincing each other, competition and many more. Many studies in social networks assume simply that there is a link or not (0-1 approach) (eg. Kleinberg, 2000; Durrington, Repman & Valente, 2000; Bonacich & Lloyd, 2004). However, links may also differ in power (eg. Burt, 2004). It is also

possible to distinguish between one-way and reciprocated relations (eg. Bonacich & Lloyd, 2004; Carley & Krackhardt, 1996).

Describing the network

Sociograms and matrices

The simplest description of the network is to list and describe actors (eg. family members or not) and relations between them (eg. frequency and/or type of contacts) in order to detect their most frequent characteristics. However, such a description can be hardly called social network analysis because it does not involve any information about the structure of contacts (Allen, 2000).

Social network data are therefore usually presented in two ways: either as a sociogram or as a matrix. The sociogram is a graph consisting of nodes (depicting actors) that are connected with links (depicting relations). It is the most intuitive form of presenting network data. It provides hollistic overview and might be useful in forming hypotheses in regard to the network's parameters.

The matrix' rows and columns note the same set of actors/cases/subjects whereas the cells contain information about relations between subjects. This format of data allows for great variety of sophisticated analytical tools that help to capture specific features of the network (Hanneman, 2001).

The most typical structural parameters of a network capture: distance between nodes (how many nodes you need to go through to reach any other specific node); density of ties (how many links are present in proportion to all possible connections between nodes); size of the network (the number of nodes); centrality (number of ties to/from any given node); existence of cliques and sub-groups and many others (Hanneman, 2001; Allen, 2000; Bonacich & Lloyd, 2004).

However, most studies on social networks deal with data collected only once for the given population and therefore capturing only static phenomena. Only a few papers deal with the issue of evolution of the network's structure (Suitor, Wellman, & Morgan, 1997; Morgan, Neal, & Carder, 1996; Suitor & Keeton, 1997; Wellman et al., 1997; Ruan et al., 1997; Feld, 1997). The evolution is meant here as changes in number and quality of nodes and links. Longitudinal studies are able to capture disappearing and appearing nodes and links but it usually adds up heavily to the complexity of the model. In case of CAVES it would be tempting to focus on historic evolution of social networks in context of past shock. Unfortunately there is no historical data on social networks and it would be extremely difficult or impossible to re-contruct the structure of past networks relying on contemporary field data – due to the fact that many people must have moved out of the area or died, it would be impossible to question all (or even nearly all) members of a network, and even if it would, it must be remembered that this sort of information is very much prone do memory distortions.

Procedure – how to collect data. Overwiev

The bibliography on social networks offers examples of many various attempts to collect data about social networks. The most popular methods are in general interview-based. Some researchers ask people to list all their (relevant) contacts by name (Litwin & Landau, 2000; Marsden, 2003; Suh et al., 1997; The General Social Survey) in order to obtain information about number of contacts and some basic information (eg. sex, kind of relation). Sometimes the interview is broader, including more detailed questions about the history and quality of contacts (Curtis et al., 1995; Friedman et al., 1998; Glass et al., 1997; Lai, 2001; Latkin et al., 1995; Morgan et al., 1996; Ruan et al., 1997; Suitor & Keeton, 1997; Wellman et al., 1997). In order to better quantify the results (although taking the risk of loosing some unexpected yet important data) one may ask people to list their contacts and evaluate them on pre-defined scales (questionnaire) (Casciaro, 1998; Ferligoj & Hlebec, 1999; Feld, 1997; Jansson, 1997; Litwin, 1995). It is also possible to obtain graphic rather than verbal data by asking people to diagram their social contacts (Morgan et al., 1996; Samuelsson, 1997). All these methods may provide data that allow to reconstruct the actual structure of the network by matching people who indicated each other as their contacts. Of course in order to to be able to create such a diagram, the answers must allow for unambiguous identification of actors and the people questioned should belong to the population creating the network.

Other methods do not focus on re-creating the structure of the network, narrowing down to estimation of some important parameters, eg. its average density. Data sufficient to answer such a question may be collected by asking people how many contacts they have (Allen, 2000), showing a list common first names and ask if the person knows someone of this name (Brewer, 1997; McCarty et al., 1997) or performing a random walk (asking one person to list his/her contacts, randomly select one of them and go to the selected person (4 subsequent persons)) (Liebow et al., 1995; McGrady et al., 1995).

It is important to note that whichever way the data is obtained, it is always innacurate to some extent. In general, reports about social contacts differ significantly from data about actual frequency of such contacts as derived from observation. What is more, if an agent reports a reciprocated relation with another person, it does not neccesarily mean that the other person will report the same link. In other words, subjective reports of different actors do not always match the reports of others. The problem of accuracy of perception and report of one's social network is however rarely taken into account (Casciaro, 1998; Feld & Carter, 2002).

Social networks in CAVES

In order to create a model of the social network in the selected area of Odra River Valley there were several important steps identified:

- Selecting a method of collecting data
- Defining who will be the nodes
- Defining which links are of interest
- Creating a tool (questionnaire)
- Selecting a sample
- Questioning people
- Creating a model
- Analysing the structure of social networks.

Ideally, one should collect data from all agents forming the network of interest. In case of Odra River Valley it is however impossible – the research area consists of almost twenty municipalities inhabited by tens of thousands of people. Therefore, it has been decided that the social network of interest will be estimated from smaller sample of networks derived from field data. Since the focus of the project is on land use, which is mostly decided at the level of a household, nodes of the network represent households. Farmers were interviewed in two villages in two different municipalities balanced for their various land-use and social characteristics.

Most works on social networks focus on one type of links, eg. asking people to list all their friends or family members. As a result there is usually one model of a social network capturing social contacts in general. As pointed out by Crowell (2004), the network may look different and provide different incentives to agents if we consider variuos types of links, eg. one type of contacts may provide emotional support and protect from effects of stress, whereas another network offers access to material or informational support (Allen, 2000; Crowell, 2004).

Since it is difficult to predict which kind of contacts will have the most influence of farmers' decisions about land use, it seems reasonable to consider links of several types. Namely, it may be important to know with whom farmers create bonds by exchanging information, helping each other, making business or simply socialising. The questions asked will refer to categories broader than land-use only because it should be taken into consideration that farmers may eg. follow someone's example based on the general impression or general experience with the person rather than on extensive analysis of data related only to land-use. In other words, we take into consideration that decision making is not an entirely rational process but does involve heuristics, such as following the example of respected others.

One or two persons from each household (the head of the family and his wife) were asked to list names of their contacts of selected kinds. In order to estimate the strenght of links they were asked how often given contacts occur (other questions, eg. about the estimated strenght of the contact would be more prone to differences in subjective estimation of the strenght. They could also be considered as more intimate and therefore not answered).

Sampling

There were two villages selected as the field of research, based on their social and landscape characteristics. The selection has been made based on opinion of key local experts and GIS experts. One village was known for multiple activities of its inhabitants, another one was selected as its inhabitants are known to be rather passive in regard to common social activities. All inhabitants of the selected villages were invited to participate. They were asked to agree on participation.

Hypotheses

Collected data is expected to allow for re-constructing the structure of the existing real network of social contacts in order to use it for modelling.

As an additional task, it will be examined whether any characteristics of the network are related to the pattern of civic activity in the selected villages. It is expected that networks in active versus non-active communities will differ in density, distribution and/or strenght of links.

Method

Measures

Data about the level of actual activity of inhabitants was collected from each respondent – the "social activity" index was calculated as a sum of positive answers to a number of questions regarding participation in various forms of social involvement (organizing festivities, protests, participation in demonstrations, volunteering etc). Another index was calculated for participation in various organizations (local committees, political parties, trade unions, parents board at school, religious associations, clubs, voluntary fire-brigades etc). Those two indexed were summed up to obtain a general index of social involvement.

Statistical analysis revealed that two selected villages indeed differed significantly in regard to involvement in social activities $t(40)=2,03$, $p<0.05$ in the predicted direction. The mean index for households in the "active" village was $M_{act}=4,22$ ($SD_{act}=3,10$) as compared to the "non-active" village

$M_{n\text{-}act}$=2,47 ($SD_{n\text{-}act}$=2,29) There was no significant differences in indexes of organization involvement.

Participants

Questionnaires were carried out with 37 respondents in the "active" village and 29 respondents in the "non-active village" that represented respectively 23 and 19 households. Interviewers were instructed to visit each household in each of the two villages and invite the couple (or a person) who was declared to be in power of decisions in the given household to participate in the study. The refusal rates were however high – data has been obtained only from 55% households in the "active" village and 36% households in the "non-active village".

Questionnaire

The questionnaire consisted of three parts – the first one contained questions regarding demographic data (educational level, gender etc). The second part was listing social activities and social organizations. Respondents were ask whether or not they participated in any event or organization, respectively. The indexes of social and organizational involvement were constructed by adding up the number of positive answers.

The third part comprised four questions about various kinds of social contacts. Respondents were asked to recall names of people they contact for different reasons (seeking information, helping each other, socializing and doing business). For each relation it had to be then specified how often such contacts took place. Responses were coded on an ordinal scale: very rare – rare – moderate – frequent – very frequent.

This data was used to reconstruct networks of relations between households.

Summary of results

Fieldwork data gathered for the CAVES project was used to create a model of social networks in Odra River Valley. Two case study villages were selected based on their social and geographical characteristics. One of the villages was expected to have a dense network of relations, as it was selected by local experts as showing high participation rate in collective actions. The other village was selected to show opposite characteristics.

In each village interviewers were instructed to visit every household and interview the couple (or person) who was declared to be in power of decision making in the given household. Respondents were asked about four types of interpersonal relations: seeking for information, mutual help, socializing

and business contacts. As a result, eight models of social networks were obtained. However, collected data is highly incomplete as refusal rates were higher than expected. Therefore, created networks cannot be considered as complete and representative models of social networks in the examined area. Analysis of networks' structure was therefore only limited to simple characteristics that could be estimated with reasonable accuracy, based on data available. Further analysis of structural parameters would be very prone to distortions due to incomplete data and therefore was not performed.

Comparison of between villages revealed that the village selected as active, based on local experts' opinion indeed scored higher on a scale of collective social activities. Also, as expected, the networks of information flow and socializing were denser in the "active" village (respondents reported more relations). There was no difference between networks of business relations and differences between help networks was only approaching the level of statistical difference.

The networks in the "active" village also contained more reciprocated relations (i.e. relations between two households that were declared to exist by both parties) – approaching the level of 20% of all relations, as compared to reciprocity indexes of the "non-active" village that were mostly below 10%.

In case of networks in the "active" village, there were a few nodes identified that served as main centers – having the biggest numbers of in-coming and out-going ties. The pattern of one leader and two-three other main actors was replicated in all four networks in this village.

Networks obtained for the "non-active" village were in all cases smaller and much less interconnected that for the "active" village. Some of them were actually so fragmented and dispersed that it can be questioned whether they constituted networks at all. It may however be due to lower response rate and therefore incomplete data.

It seems that although there were nodes that scored quite high in terms of numbers of out-going and in-coming ties, there was no apparent leader, dominating all the scales. In various networks there were different leaders.

It should be mentioned that respondents often reported having contacts with people living outside the village. Such persons were never mentioned by more than one respondent. It may be thus concluded that external ties of villages are non-redundant but on the other hand they are prone to changes or losses.

References

Allen, W.D. (2000). Social networks and self-employment. *Journal of Socio-Economics, 29,* 487-501.

Baker, W.E., & Faulkner, R.R. (2004). Social networks and loss of capital. *Social Networks, 26,* 91-111.

Bonacich, P. & Lloyd, P. (2004). Calculating status with negative relations. *Social Networks, 26,* 331-338.

Brewer, D.D. (1997). No associative biases in the first name cued recall procedure for eliciting personal networks. *Social Networks, 19,* 345-353.

Burt, R.S. (2004). Structural holes and good ideas. *American Journal of Sociology, 110(2),* 349-399.

Carley, K.M., & Krackhardt, D. (1996). Cognitive inconsistencies and non-symmetric friendship. *Social Networks, 18,* 1-27.

Casciaro, T. (1998). Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Social Networks, 20,* 331-351.

Crowell, L.F. (2004). Weak ties: a mechanism for helping women expand their social networks and increase their capital. *The Social Science Journal, 41,* 15-28.

Curtis, R., Friedman, S.R., Neaigus, A., Jose, B., Goldstein, M., & Ildefonso, G. (1995). Street-level drug markets: Network structure and HIV risk. *Social Networks, 17,* 229-249.

Durrington, V.A., Repman, J., & Valente, T.W. (2000). Using social network analysis to examine the time of adoption of computer-related services among university faculty. *Journal of Research on Computing in Education, 33(1),* 16-27.

Faust, K., Entwisle, B., Rindfuss R.R., Walsh, S.J., & Sawangdee, Y. (1999). Spatial arrangement of social and economic networks among villages in Nang Rong District, Thailand. *Social Networks, 21,* 311-337.

Feld, S.L. (1997). Structural embeddedness and stability of interpersonal relations. *Social Networks, 19,* 91-95.

Feld, S.L., & Carter, W.C. (2002). Detecting measurment bias in respondent reports of personal networks. *Social Networks, 24,* 365-383.

Ferligoj, A., & Hlebec, V. (1999). Evaluation of social network measurement instruments. *Social Networks, 21,* 111-130.

Friedman, S.R., Neaigus, A., Jose, B., Curtis, R., & Des Jarlais, D. (1998). Networks and HIV risk: An introduction to social network analysis for harm reductionists. *International Journal of Drug Policy, 9,* 461-469.

Glass, T.A., Mendes de Leon, C.F., Seeman, T.E., & Berkman, L.F. (1997). Beyond single indicators of social networks: A LISREL analysis of social ties among the elderly. *Social Science & Medicine, 44(10),* 1503-1517.

Hanneman, R.A. (2001). *Introduction to social network methods.* Textbook available on-line, retrieved June 02, 2005 from http://faculty.ucr.edu/%7Ehanneman/SOC157/Software/Downloadindex.html

Jansson, I. (1997). Clique structure in school class data. *Social Networks, 19,* 285-301.

Kleinberg, J. (2000). Navigation in a small world. *Nature, 406,* 845.

Lai, G. (2001). Social support networks in urban Shanghai. *Social Networks, 23,* 73-85.

Latkin, C., Mandell, W., Oziemkowska, M., Celentano, D., Vlahov, D., Ensminger, M., & Knowlton, A. (1995). Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. *Drug and Alcohol Dependence, 38,* 1-9.

Liebow, E., McGrady, G., Branch, K., Vera, M., Klovdahl, A., Lovely, R., Mueller, C., & Mann, E. (1995). Eliciting social network data and ecological model-building: Focus on choice of name generators and administration of random-walk study procedures. *Social Networks, 17,* 257-272.

Litwin, H. (1995). The social networks of elderly immigrants: An analytic typology. *Journal of Aging Studies, 9(2),* 155-174.

Litwin, H., & Landau, R. (2000). Social network type and social support among the old-old. *Journal of Aging Studies, 14(2),* 213-228.

Marsden, P.V. (2003). Interviewer effects in measuring network size using a single name generator. *Social Networks, 25,* 1-16.

McCarty, C., Bernard, H.R., Killworth, P.D., Shelley, G.A., & Johnsen, E.C. (1997). Eliciting representative samples of personal networks. *Social Networks, 19,* 303-323.

McGrady, G., Marrow, C., Myers, G., Daniels, M., Vera, M., Mueller, C., Liebow, E., Klovdahl, A., & Lovely, R.(1995). A note on implementation of random-walk design to study adolescent social networks. *Social Networks, 17,* 251-255.

Morgan, D.L., Neal, M.B., & Carder, P. (1996). The stability of core and peripheral networks over time. *Social Networks, 19,* 9-25.

Ruan, D., Freeman, L.C., Dai, X., Pan, Y., & Zhang, W. (1997). On the changing structure of social networks in urban China. *Social Networks, 19,* 75-89.

Samuelsson, M.A.K. (1997). Social networks of children in single-parent families: Differences according to sex, age, socioeconomic status and housing-type and their associations with behavioural disturbances. *Social Networks, 19,* 113-127.

Scott, J. (2000). Handling relational data. *Social Network Analysis: A Handbook*. London: Sage.

Suh, T., Mandell, W., Latkin, C., & Kim, J. (1997). Social network characteristics and injecting HIV-risk behaviors among street injection drug users. *Drug and Alcohol Dependence, 47,* 137-143.

Suitor, J., & Keeton, S. (1997). Once a friend, always a friend? Effects of homophily on women's support networks across decade. *Social Networks, 19,* 51-62.

Suitor, J., Wellman, B., & Morgan, D.L. (1997). It's about time: How, why, and when networks change. *Social Networks, 19,* 1-7.

Wellman, B., Yuk-lin Wong, R., Tindall, D., & Nazer, N. (1997). A decade of network change: Turnover, persistence and stability in personal communities. *Social Networks, 19,* 27-50.