

Recall Versus Recognition: Comparison of the Two Alternative Procedures for Collecting Social Network Data

Valentina Hlebec¹

Abstract

In the paper the results of an experiment in measuring the effect of two alternative methods for collecting social network data are presented. Recall and recognition of the communication flow, identified between twelve members and advisers of the Student Government of the University in Ljubljana, were compared according to:

- the size of egocentric networks and
- the stability of naming.

The hypotheses were:

- the average size of the recalled egocentric network would be smaller than the recognized one and the differences would be minor,
- the respondent with larger recalled network would have larger recognized network.

All hypotheses were confirmed for one of the three defined relations. Difficulties with the two other relations could be explained by the two different content criteria included in the questions for identifying these two relations.

Keywords: Social network analysis; Measurement; Data collection procedures; Communication flow.

1 Introduction

Social network analysis is a recent research field which is more and more used in the social sciences. The main advantage of the social network analysis is the ability for detection, measurement and analysis of social relations. Some of the researchers using social network analysis are concerned with basic measurement problems connected with data collection. Holland and Leinhardt (1973) examined the effect of the fixed choice procedure for collecting social network data, Granovetter (1982)

¹Faculty of Social Sciences, University of Ljubljana, Kardeljeva pl. 5, 61 109 Ljubljana, Slovenia

proved that it is necessary to collect also weak ties not only strong ties, Laumann et al (1983) were concerned with boundary specification problem in collection of network data, Sudman (1985, 1988) described basic procedures for measuring the size of an egocentric network, Hammer (1984), Bien et al (1991) and others also paid some attention to these problems.

Network data can be collected on several different ways. Self-reports (face to face, telephone and mail interview), archival records, direct observation, competent informant, diary and electronic communication media can be found in literature as a data source. The usual way of collecting social network data is face to face interview. Our experiment was constructed with the purpose to examine the effects of the alternative procedures for collecting social network data with face to face interview.

Alternative procedures (recall, recognition and numerical estimation) for measuring the size of the egocentric networks (Sudman 1985, 1988) and procedures for measuring the strength of the emotional relation (Holland and Leinhardt 1973) were adjusted to the identification of the social relations (communication flow) among the members of the complete network for the purposes of our experiment.

2 Definition of a network

Let $E = \{X_1, X_2, \dots, X_n\}$ be a finite set of *units*. The units are related by one (or more) binary (or numerical) *relations*

$$R_t \subseteq E \times E, \quad t = 1, \dots, r$$

which determine a *network* (Batagelj, Ferligoj, and Doreian 1992)

$$\mathcal{N} = (E, R_1, R_2, \dots, R_r)$$

A single relation R can be described by a corresponding binary matrix $\mathbf{R} = [r_{ij}]_{n \times n}$ where

$$r_{ij} = \begin{cases} 1 & X_i R X_j \\ 0 & \text{otherwise} \end{cases}$$

In some applications r_{ij} can be a nonnegative real number expressing the strength of the relation R between units X_i and X_j .

Relations among the units can also be presented by a graph, called a sociogram. In a graph, units are represented by a set of points and a set of relations are represented by lines, drawn between pairs having direct connections.

3 Measurement methods

In the experiment two alternative methods were used for collection of network data:

1. Recall: Members of the group were asked to identify the members of their egocentric networks by memory. The criteria for enumeration was frequency of the recalled communications.

2. **Recognition:** The list of all members of the group was given to each member. They were first asked to identify who they communicate with and than to select the persons they communicate with most often.

The number of listed persons was not limited in any method.

4 Hypotheses

According to the results of previous similar analyses two hypotheses were the starting point of this study:

1. The average size of the recalled egocentric network would be smaller than the recognized one and the differences would be minor,
2. The respondent with larger recalled network would have larger recognized network.

These hypotheses were tested by the experiment described in the following sections.

5 Data collection

The analyzed network consisted of communication interactions among twelve members and advisors of the Student Government at the University in Ljubljana. The results of the measurement are not real interactions among actors but cognition about communication interactions. Data were collected with face to face interviews which lasted from 20 to 30 minutes. Interviews were conducted in May 1992.

Communication flow among actors was identified by three questions:

1. Who of the members and advisors of the Student government do you (most often) informally discuss with?
2. Which members and advisors of the Student Government do you (most often) ask for an opinion?
3. Which of the members and advisors of the Student Government (most often) ask you for an opinion?

The content of the communication flow was limited to the matters of the Student Government. The time frame was also defined: questions were referred to the six months period (from the formation of the government to the day of the interview). The obtained relations are represented by matrices in Appendix. ¹

One respondent refused to cooperate in the experiment. As he was not considered in the analysis, the network consists of eleven actors.

¹Only one respondent listed all the others under the recognition method for two relations (discussion, asking for an opinion). For that respondent the first group was defined arbitrary. The cut point was determined by the average number of selected persons at the recognition method (3.5 for the first relation and 4.5 for the second).

6 Results

The average size of the egocentric network (the average number of persons named by each respondent), standard deviation, minimum and maximum of each measured network are given in the following table:

	mean	st.d.	min	max
R1 (A)	3.73	1.42	1	6
(B)	4.36	1.63	2	8
R2 (A)	2.62	1.21	1	5
(B)	3.27	1.42	2	7
R3 (A)	3.73	2.45	0	8
(B)	3.82	1.49	1	6

Notations in the table mean: R1 (discussion), R2 (asking for an opinion), R3 (being asked for an opinion), A (recall), B (recognition).

From the table we can see that the first hypothesis was confirmed for all three relations. The average size of the recalled egocentric network was smaller than the recognized one. The number of the recognized persons was on average larger than the number of the recalled persons. The maximum number of persons named was higher for recognition for all three relation. The standard deviations were also higher for recognition for the first two relations (discussion, asking for an opinion).

The differences between two methods were analyzed according to Sudman (1985). They were computed on the basis of comparison of individual responses. The differences between an individual's response to recognition and recall was squared and summed to give the absolute differences for a specified relation. The relative differences were obtained by dividing the absolute differences by the average size of the recognized egocentric network for a specified relation. Absolute and relative differences are:

	ABS	REL
R1	1.00	0.22
R2	4.10	0.62
R3	1.07	0.34

The differences were smaller for discussion (R1) and being asked for an opinion (R3). The differences are comparable with Sudman's results for a small working group (18 members) a and social club (19 members) only for those two relations. The hypothesis can be confirmed only for the first and third relation.

Table 1: The size of the egocentric networks

UNIT	R1		R2		R3	
	A	B	A	B	A	B
1	3	4	2	2	4	6
2	1	3	2	2	8	6
3	6	8	5	3	2	1
4	2	2	1	2	6	5
5	5	6	2	4	4	3
6	5	5	1	7	2	4
7	4	4	3	3	4	5
8	4	4	3	3	7	6
9	4	5	3	3	0	1
10	4	4	4	4	2	3
11	3	3	3	3	2	2

If we assume, that the difference in the size of the recalled and the recognized egocentric network (Table 1) is due to different types of measurement, than we expect, that the respondent with larger recalled network would have larger recognized network. The relationship between the size of the recalled network and the recognized network was analyzed by Pearson's correlation coefficient.

Pearson's correlation coefficient between the size of the recalled and recognized egocentric network are:

1. relation R1: $r = 0.87$,
2. relation R1: $r = -0.17$,
3. relation R2: $r = 0.83$.

The hypothesis can be confirmed for the first (discussion) and the third (being asked for an opinion) relation. The results for the second (asking for an opinion) relation are not surprising with regard to the standard deviation and relative differences which were high.

In the Table 2 the differences between naming by recognition and recall are presented. There is a plus if a person was named by recognition and not by recall and there is a minus if a person was named by recall and not by recognition. Positive differences were expected.

The negative differences were minor for the first relation (discussion), and larger for the others. The negative differences can be explained with the limitation to the first recognized group or with arbitrary determined rank for one respondent for the first relation and with two different content criteria used in two questions for identifying these relations. Each member and advisor of the Student Government has

Table 2: Stability of naming

UNIT	R1		R2		R3	
	+	-	+	-	+	-
1	1		1	1	2	
2	2					2
3	2		1	3	1	2
4			2	1	1	2
5	1		2		1	2
6			6		2	
7					2	1
8						
9	2	1	1	1	1	
10	1	1			2	1
11	1	1	1	1	1	1

a separate working domain. Each respondent can ask the others about their special working domain or about general matters. It is obvious that the two questions included at least two different content criteria. The criteria were not explicitly defined, so the respondent could answer these questions in several different ways. It was not possible to recognize which of the criteria was used for answering these questions.

7 Conclusion

All hypotheses were confirmed for one of the three defined relations. Difficulties with the other two relations can be explained with at least two different content criteria included in the questions for identifying these two relations. It was not possible to find which of the criteria was used for answering these questions.

Regardless of the methodological error in the formulation of the questions, the hypotheses can be confirmed. Recall would yield worse results and according to the small size of the group, the differences would be minor. Recognition would comprise data collected with recall and the differences in classification are mostly due to different procedures.

In the case of small number of units in network the data can be correctly collected by both procedures. Recognition is more convenient for collecting such data, because it also includes data collected with recall. The first recognized group according to some criteria of strength correspond with persons named by recall.

References

- [1] Batagelj, V., Ferligoj, A., and Doreian, P. (1992): Direct and indirect methods for structural equivalence. *Social Networks*, **14**, 63-90.
- [2] Bien, W, Marbach, J., and Neyer, F. (1991): Using egocentered networks in survey research. A methodological preview on an application of social network analysis in the area of family research. *Social Networks*, **13**, 75-90.
- [3] Granovetter, M (1982): The strength of weak ties: A network theory revisited. In: P.V. Marsden and N. Lin (Eds.): *Social Structure and Network Analysis*. Beverly Hills: Sage, 105-130.
- [4] Hammer, M. (1984): Explorations into the meaning of social network interview data. *Social Networks*, **6**, 341-371.
- [5] Holland, P.W. and Leinhardt, S. (1973): The structural implications of measurement error in sociometry. *Journal of Mathematical Sociology*, **3**, 85-111.
- [6] Laumann, E.O., Marsden, P.V., and Prensky, D. (1983): The boundary specification problem in network analysis. In: R.S. Burt, and M.J. Minor (Eds): *Applied Network Analysis: A Methodological Introduction*. Beverly Hills: Sage, 18-34.
- [7] Sudman, S. (1985): Experiments in the measurement of the size of social networks. *Social Networks*, **7**, 127-151.
- [8] Sudman, S. (1988): Experiments in measuring neighbour and relative social networks. *Social Networks*, **10**, 93-108.

Appendix: Six measured relations

A - discussion, recall	B - discussion, recognition
1 0 1 1 0 0 1 0 0 0 0 0	1 0 1 1 0 0 1 0 0 0 1 0
2 0 0 0 0 0 0 0 1 0 0 0	2 0 0 0 1 0 1 0 1 0 0 0
3 1 1 0 1 0 1 1 1 0 0 0	3 1 1 0 1 1 1 1 1 0 1 0
4 0 0 0 0 0 0 1 1 0 0 0	4 0 0 0 0 0 0 1 1 0 0 0
5 0 1 0 1 0 1 1 1 0 0 0	5 0 1 0 1 0 1 1 1 0 1 0
6 0 1 0 1 1 0 1 1 0 0 0	6 0 1 0 1 1 0 1 1 0 0 0
7 0 0 0 1 0 0 0 1 1 0 1	7 0 0 0 1 0 0 0 1 1 0 1
8 0 1 0 1 0 0 1 0 0 0 1	8 0 1 0 1 0 0 1 0 0 0 1
9 0 0 0 1 0 0 1 1 0 0 1	9 0 1 1 0 0 0 1 1 0 0 1
10 1 0 1 1 1 0 0 0 0 0 0	10 1 1 1 0 1 0 0 0 0 0 0
11 0 0 0 0 0 1 0 1 1 0 0	11 0 0 0 1 0 1 0 1 0 0 0

C - asking for opinion,
recall

1 0 1 0 0 0 1 0 0 0 0 0
 2 0 0 0 1 0 0 0 1 0 0 0
 3 1 1 0 1 0 1 0 1 0 0 0
 4 0 1 0 0 0 0 0 0 0 0 0
 5 0 1 0 0 0 1 0 0 0 0 0
 6 0 1 0 0 0 0 0 0 0 0 0
 7 0 0 0 1 0 0 0 1 1 0 0
 8 0 1 0 1 0 0 0 0 0 0 1
 9 0 0 0 0 0 0 1 1 0 0 1
 10 1 1 1 0 1 0 0 0 0 0 0
 11 0 0 1 0 0 1 0 1 0 0 0

D - asking for opinion,
recognition

1 0 1 1 0 0 0 0 0 0 0 0
 2 0 0 0 1 0 0 0 1 0 0 0
 3 1 1 0 0 0 0 0 0 0 1 0
 4 0 0 0 0 0 0 1 1 0 0 0
 5 0 1 0 0 0 1 0 1 0 1 0
 6 1 1 1 1 1 0 1 1 0 0 0
 7 0 0 0 1 0 0 0 1 1 0 0
 8 0 1 0 1 0 0 0 0 0 0 1
 9 0 1 0 0 0 0 1 1 0 0 0
 10 1 1 1 0 1 0 0 0 0 0 0
 11 0 0 0 1 0 1 0 1 0 0 0

E - being asked for opinion,
recall

1 0 1 1 0 0 0 0 0 0 1 1
 2 1 0 1 1 1 1 1 1 0 1 0
 3 0 1 0 0 1 0 0 0 0 0 0
 4 0 1 1 0 0 1 1 1 0 0 1
 5 0 1 0 0 0 1 1 0 0 1 0
 6 0 1 0 0 0 0 0 0 0 0 1
 7 0 0 0 1 0 0 0 1 1 0 1
 8 0 1 1 1 0 1 1 0 1 0 1
 9 0 0 0 0 0 0 0 0 0 0 0
 10 0 1 0 0 1 0 0 0 0 0 0
 11 0 1 0 0 0 0 0 0 1 0 0

F - being asked for opinion,
recognition

1 0 1 1 1 0 0 0 1 0 1 1
 2 1 0 1 1 0 1 0 1 0 1 0
 3 1 0 0 0 0 0 0 0 0 0 0
 4 0 0 1 0 0 1 1 1 1 0 0
 5 1 1 0 0 0 1 0 0 0 0 0
 6 1 1 0 0 1 0 0 0 0 0 1
 7 1 0 0 1 0 1 0 1 1 0 0
 8 0 1 0 1 0 1 1 0 1 0 1
 9 0 0 0 0 0 0 1 0 0 0 0
 10 1 0 1 0 1 0 0 0 0 0 0
 11 0 0 0 0 0 0 1 0 1 0 0