Triad Census Usage for Communication Network Analysis

Olga M. Zvereva

Ural Federal University, Ekaterinburg, Russia
{OM-Zvereva2008@yandex.ru}

Abstract. Small groups occupy an important place in our lives: we live in a family, we work in a structural department of a company, and we can be members of a school class or a university group. SNA-methodology helps us to answer many life's vital questions, one of which is how to predict social group sustainability and its existence for a long time. In this paper it is argued that a structurally balanced group is likely to be a sustainable one. Some structural balance theories are under consideration in order to evaluate the structural balance level in a communication network. The triad census characteristic is used as an analytical basis. It was demonstrated that the triad census is a very useful and informative characteristic, and, while its consideration and comparison with the calculated census for a random Bernoulli graph, some additional phenomena were revealed. In this research work a new method of data collection and its processing was used; the advantages of this new approach are also discussed.

Keywords: SNA-methodology, communication network, structural balance, triad census, transitivity

1 Introduction

If we hear today "a social network", we will imagine a global network, such as Facebook, Instagram, Twitter, and etc., and we almost forget that a small group is also a social network but of a smaller size. We are members of a number of small groups: we live in a family, we work in a structural department of a company, and we can be members of a school class or university group.

There are two clearly definable periods in research of small group networks: the 1950s and 1960s (the "early era"), and since1990s till our days (the "current era") [1]. While sociometric methods were very popular in the early era, SNA-methodology (SNA-Social Network Analysis) has become the main methodology in "the current era" research.

Armed with information tools, SNA combines the graph theory, mathematical statistics, and sociological methods of data collection and its analysis and interpretation.

Barry Wellman [2] was one of the first who provided the theoretical basis for network approach of social group analysis. He stated that persons' behavior is well predicted by examining not their drives, attitudes, or demographic characteristics, but rather the web of relationships in which they are embedded. He argued that the flow of information and resources between two individuals was depended not simply on their relationship to each other but on their relationships to everybody else.

Various type ties could be revealed between social group members, but we have focused on those of communication type. We tried to fix all the communications in which students took part while their education activities. These fixations were made by students themselves through their usage of the self-designed information system.

Many sociologists consider communications to be the main relations in a social group, and we paid the main attention to the communications which originated in the educational process.

N. Luhmann argued that social systems were constructed of communications. “Their (social systems) elements are communications which are recursively produced and reproduced by a network of communications and which cannot exist outside of such a network” [3].

This idea was one of the main reasons to focus on communication ties in our research. Students from 4 different university groups took part in it during the educational year.

The main issue under consideration was a group sustainability. We were interested, why some social groups were keeping their integrity, maintaining internal relations, even when formal ties were destructed, while other groups failed to exist under the similar conditions. We try to find sustainability premises in the group structure. What is the group index which could propose the evidence of its sustainability, and what concepts are to deliver the theoretical foundation for this choice?

We argue that structurally balanced group is more likely to be a sustainable one. As the main index for balance evaluation a triad census characteristic is proposed. In the next paragraphs description of the balance theories, the triad census specification, the method of data collection and this research results are delivered.

2 Balance theories

There are several theories of a structural balance. They are described in details in [4, 5, 6, 1]. For better understanding of the further research results we propose their description in brief.

One of the first theories, which gave the birth to a set of balance models, was the cognitive theory of Heider. Heider’s theory [7] posited that there were a number of psychical forces in the individual cognitive field which were oriented towards the balance preservation. According to Heider, a balance is not a real ratio of element forces but its perception by an individual. If an individual considered another individual to support positive relations, then any negative act would “spoil” the whole picture, and psychical forces would try to reestablish the equilibrium.

From the behavioral point of view, the important consequence of Heider’s theory is the assumption that any positive relation is a transitive one (e.g. “a friend of my friend is also my friend” is true), and, vice versa, any negative relation is an intransitive one (e.g. “an enemy of my enemy is my enemy” is false). If these conditions are satisfied, the structural balance exists.

This model, mathematically expressed by Harary and Cartwright [8], was extended to consider a graph of any size. This concept was used for social networks, as a network was presented as a signed graph with edges having positive or negative valence.

Harary-Cartwright’s structure theorem confirms that “a s-graph (signed graph) is the balanced one if and only if its points can be separated into two mutually exclusive subsets such that each positive line joins two points of the same subset and each negative line join points from different subsets” [8].

In another words, a group presented by a balanced graph can be partitioned into two antagonistic subgroups (one of these subgroups can be empty) so, that internal subgroup links must be positive, and external subgroup links must be negative. According to E.C. Johnsen [9], in a directed graph mutual ties are treated as positive and null ties as negative, thus, a balanced directed graph must have only mutual ties within subgraphs and only null ties between subgraphs.

It is difficult to find such partition in reality, that is why, new empirical models were proposed. Davis [10] extended the structural balance to the more sociologically reasonable notion of clusterability, which allows existence of more than two subgroups. But the rules were similar to the previous model: positive (mutual) links in a subgroup (subgraph), and negative (null) links between subgroups (subgraphs).

In all previously discussed models, only mutual (symmetric) and null ties are permitted. In reality, we often have the symmetric kind of relationships. Most of communication ties (such as who talks to whom, who gives information or advice to whom), affective ties (such as who likes whom, or who trusts whom) or formal ties (such as who reports to whom) are asymmetric in their nature.

The Ranked Clusters Model assumes that any group is a subgroup hierarchy where every hierarchical level has at least one subgroup. This model extends clustering with permission of directed (asymmetric) ties between subgroups, with orientation of the directed ties consistent with hierarchical ordering in which ties are directed from ”lower” to “higher” levels (any lower level subgroup member prefers those who are members of the higher level subgroups). Davis and Leinhardt are considered to be the authors of this model [11].

The Transitivity Model is the most general model of all discussed and subsumes the other models as special cases. The main rule in this model is as follows: if there are 3 members named A, B, C in a group, and A has a tie with B, B has a tie with C, then .A must have a tie with C.

How can we use these theories and reveal the tructural balance existence in a graph (and a social group also)? It is reasonable to analyze triadic census for this purpose.

3 Dyads, triads and triad census

The minimal structural element in a social network is a dyad – two actors who are probably tied with each other. But most scholars argue that the minimal social group is a triad. Arguments pro this concept were delivered by D. Krackhardt in [12].

Referring to Simmel, who was the first to state the fundamental difference between a dyad and a triad, D.Krachardt clearly expressed three grounds for this difference existence. The first ground is that a dyad preserves the individuality of both dyad members in comparison with a triad, where any member's interests could be suppressed by the other group members for the interests of this group as a whole. The second ground is that in a triad every member has less bargaining power than in a dyad, because of the less threat of the group destruction, when one member leaves this triad. The third ground is that any conflict, which is inevitable in any relationship over time, is more readily managed and resolved in a triad.

Holland and Leinhardt's [13] proposition was, that many important theories about social relations can be tested by means of hypotheses about the triad census in a social network. They focused on directed rather than non-directed graphs. Triads in a directed graph might belong to one of the sixteen isomorphism classes, as presented in Fig. 1. This figure uses the standard MAN labels indicating the number of mutual, asymmetric, and null dyads in a triad, along with an additional letter for direction (U, D, C, or T) when necessary.

The triad census for a network is summarized in a 16 element vector ($\vec{T}$):

$\vec{T} = (t_{003}, t_{012}, t_{102}, t_{021D}, t_{021U}, \ t_{021C}, t_{111D}, t_{111U}, t_{030T}, t_{030C}, t_{201}, t_{120D}, t_{120U}, t_{120C}, t_{210}, t_{300})$,

where t_{type} is the number (or frequency) of the "*type*" triads (of the appropriate isomorphism class) in a network.

This characteristic may be considered to be rather complex, but it definitely determines the network graph. It implicitly includes indexes of network reciprocity, clustering, density, and it reflects the structural balance level in a network graph.

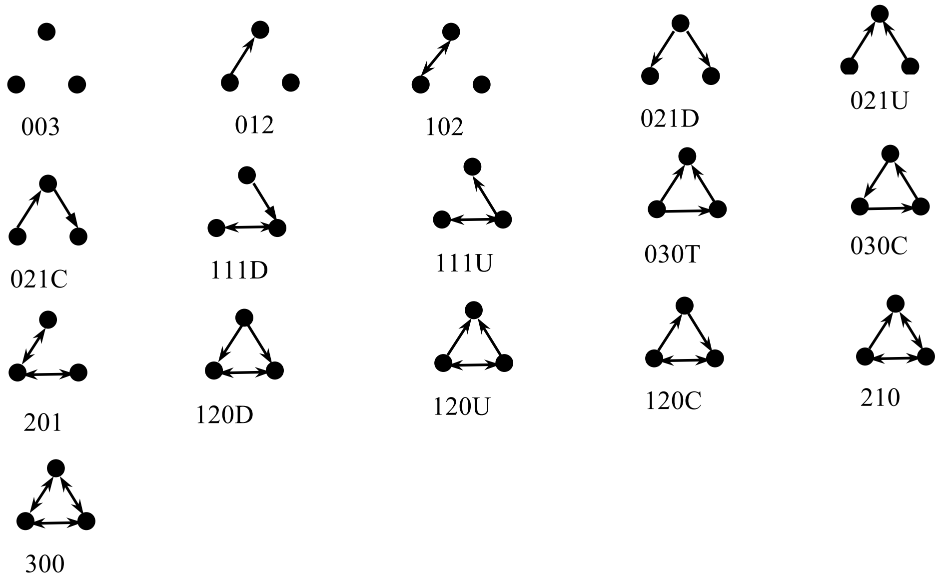

Fig. 1. Triad Census

In Tab.1 discussed theories of structural balance and permitted triad types for every theory are collected.

Table 1. Structural Balance Theories and Permitted Triad Types

№	Theory (model)	Authors	Permitted Triad Types
1	Theory of Cognitive Balance	F.Heider, D.Cartwright, F.Harary	300, 102
2	Clustering Model	J.Davis	300, 102, 003
3	Ranked Clusters Model	J.Davis,S. Leinhardt	300, 102, 003, 120D, 120U, 030T, 021D, 021U
4	Transitivity Model	P.W.Holland, S.Leinhardt	300, 102, 003, 120D, 120U, 030T, 021D, 021U, 012, 210.

In [14, 15] it is proved the 210 triad type can be considered as permitted for the Transitivity model as it has the degree of transitivity of 0.75. Thus, we argued that this triad type must be included in the list of permitted triad types for the Transitivity Model Concept.

4 Data collecting methods

Interviews, questionnaires or observation are the most common survey methods for data collecting. Each of them has its own advantages and disadvantages and might be used under the certain conditions [16].

In this research we used the different approach: data was collected while "KOMPAS TQM" system implementation was processed. This information system was designed on the basis of ideas and concepts proposed by G. Vodyanov [17]. It was engineered by students and professionals of Ural Federal University and implemented in one of its economic educational departments. It is a quality management system and is geared towards customer feedback reception. But from sociometric point of view this collected data could be interpreted as communications evaluating data, as students tried to evaluate services delivered by different departments and enterprises, and their colleagues as well.

This system supports the process of regular communication result evaluation. System users enter positive or negative marks from the certain range. Every mark must be followed by a comment. Thus, the marks entered into the system reflect the real communications between system users and they are confirmed by comments. The mark sign ("+" or "-") characterizes the "information receiver" attitude to the communication result in a whole (positive or negative) while numerical value (from 1 to 5 points) reflects the strength (weight) of this communication (i.e. the usefulness degree for receiver).

Data was collected for the period of two semesters of the 2013-2014 educational year. Students, having been registered in the system in advance, regularly input their marks, thereby fixing and evaluating their communications.

The engineered system has web interface, and every user can work with it from his/her home computer or even from his/her smartphone.

The kernel of the information system is a data base for collected data storing. Two data base sets were formed: the first one with data of winter semester and the second – for the summer semester. The "winter data set " consists of 4846 records and the "summer" one has 4769 records (it means that 4846 marks were inputted into winter data base, and 4769 – into summer data base). If we take in account only marks related to personal communications (exclude marks which evaluate the results of interaction with different organizations), the data bases include 2805 and 3394 records accordingly.

This method of data collecting has real advantages over traditional methods. First, it delivers data which is more objective, because it is a sum result of several marks not the result of a momentary reaction, which can depend on the respondent's mood or state of health.

Second, there is a time interval between a communication act and the mark input, a respondent has enough time for thinking and real evaluating of this communication.

Third, we can elicit different type data from data bases and form different sociomatrices for future research. One data set might include the sum marks put by every respondent to the other respondents, and the other set might be consisted of mark counts. The second set can be very useful for embedded ties analysis, one of the main issues under consideration in economical sociology [19].

5 Data for analysis

Four groups of students were in the focus of this research: two groups of the last year of education (identified as 34^{th} and 35^{th}) and two groups of freshmen (identified as 44^{th} and 45^{th}). For analysis binary data was received: every group matrix with mark counts was dichotomized. The rule was as follows: if a matrix element is greater than zero, it was converted into 1, and set equal to zero otherwise. Every social network of a student group is presented by a binary directed graph in this case. Graphs visualizing communication networks of the 35^{th} and 44^{th} groups are shown in Fig.2 (males are shown as blue squares and females as pink circles). It is distinct that the 35^{th} group network is denser than the network of the 44^{th} group. It is proved by the indices which are calculated.

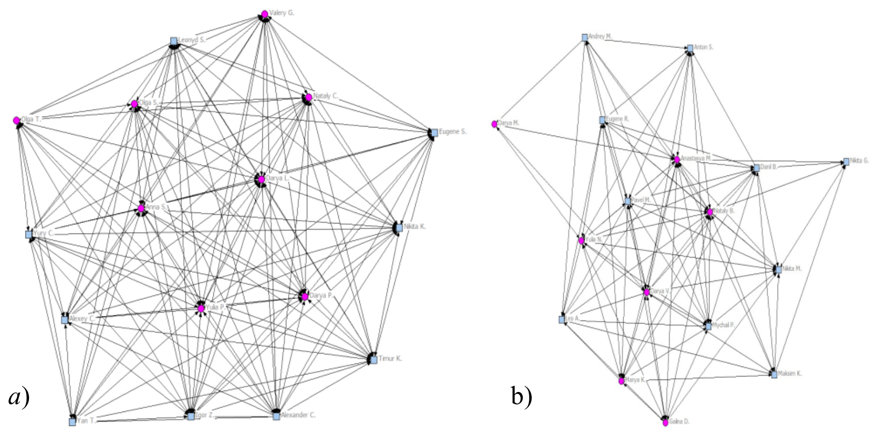

Fig. 2. Communication Network Graphs of the 35th (a) and of 44th (b) Groups

Several group characteristics are presented in Tab.2, some of them are rather clear and it is not necessary to discuss them.

Density is one of the main network characteristics (it is equal to the proportion of observed to possible edges). For directed graphs, the density is calculated as $L/n(n-1)$, where L is the number of arcs in the observed graph, and n is the number of vertices [20]. The density values (Δ)will be necessary for the further research.

Table 2. Group Characteristics

Group Id.	Year of Education	Number of Students	Males	Females	Network Density (Δ)	Clustering Coeff.	Reciprocity Coeff
34	4	13	6	7	0.66	0.715	0.515
35	4	17	9	8	0.809	0.822	0.719
44	1	17	10	7	0.250	0.616	0.528
45	1	30	15	15	0.201	0.371	0.336

As one can see from Tab.2 groups are comparable in their dimensions, the 45th group is larger than others, and the 34th group has the minimal number of members, but every group might be considered as a small group.

As for the densities, it appears that groups from the first year of education have less density values than those from the fourth year. It is a predictable fact: students from the first year groups have less time to set relations with each other and communicate rarely. The maximum density value is in the 35th group.

6 Triad Census Analysis

For every group triad census was engineered with the help of UCINET 6.0 for Windows, and the necessary percentage values were calculated with the help of Excel. All results are collected in Tab. 3.

In SNA-methodology comparison with random graph characteristics is a very popular and really useful method. This approach reveals how the social nature of a graph influences on its structure. We has followed this tradition and calculated the triad censuses for Bernoulli random digraphs of the same densities (they are presented in Table 3 as Theoretical data).

Mathematical formulas for triad type probability calculations are proposed in [21] and are shown in Table 3 (column 2). All values are presented in the percentage form.

Table 3. Triad Censuses for Groups

Triad Type	Formula	34 Group		35 Group		44 Group		45 Group	
		Theor. (%)	Real (%)	Theor. (%)	Real (%)	Theor. (%)	Real (%)	Theor. (%)	Real (%)
003	$(1-\Delta)^6$	0.15	1.05	0.00	0.00	17.80	29.26	26.02	38.00
012	$6\Delta(1-\Delta)^5$	1.80	2.10	0.12	0.29	35.60	23.97	39.27	25.94
102	$3\Delta^2(1-\Delta)^4$	1.75	2.45	0.26	1.62	5.93	12.79	4.94	11.95
021D	$3\Delta^2(1-\Delta)^4$	1.75	9.09	0.26	0.74	5.93	6.47	4.94	5.94
021U	$3\Delta^2(1-\Delta)^4$	1.75	1.40	0.26	0.59	5.93	3.24	4.94	2.44
021C	$6\Delta^2(1-\Delta)^4$	3.49	0.00	0.52	0.15	11.87	1.91	9.88	1.06
111D	$6\Delta^3(1-\Delta)^3$	6.78	0.35	2.21	2.21	3.96	0.59	2.49	0.96
111U	$6\Delta^3(1-\Delta)^3$	6.78	12.59	2.21	4.26	3.96	11.91	2.49	7.04
030T	$6\Delta^3(1-\Delta)^3$	6.78	6.99	2.21	1.32	3.96	1.03	2.49	1.01
030C	$2\Delta^3(1-\Delta)^3$	2.26	0.00	0.74	0.15	1.32	0.00	0.83	0.00
201	$3\Delta^4(1-\Delta)^2$	6.58	2.80	4.69	5.88	0.66	4.26	0.31	2.44
120D	$3\Delta^4(1-\Delta)^2$	6.58	2.80	4.69	4.85	0.66	0.00	0.31	0.07
120U	$3\Delta^4(1-\Delta)^2$	6.58	23.08	4.69	8.09	0.66	2.06	0.31	1.50
120C	$6\Delta^4(1-\Delta)^2$	13.16	0.35	9.38	2.94	1.32	0.15	0.63	0.07
210	$6\Delta^5(1-\Delta)$	25.55	17.48	39.71	33.53	0.44	1.18	0.16	0.86
300	Δ^6	8.27	17.48	28.03	33.38	0.02	1.18	0.01	0.71

Analyzing data from Table 3 one can make the following conclusions:

1. The less density value we have in a network, the more triads in this network are of the types from the beginning of a triad census, and, vice versa, the denser networks are constructed of triads which types are at the end of a triad census. For the 34^{th} and 35^{th} groups the most "popular" triads are of 210 and 300 types and for the groups with small density the most "popular" triad types are 003 and 012. In Fig. 3 one can see triad quantity distribution for real groups, and in Fig.4 we propose the same kind of distribution for Bernoulli graphs of various densities.

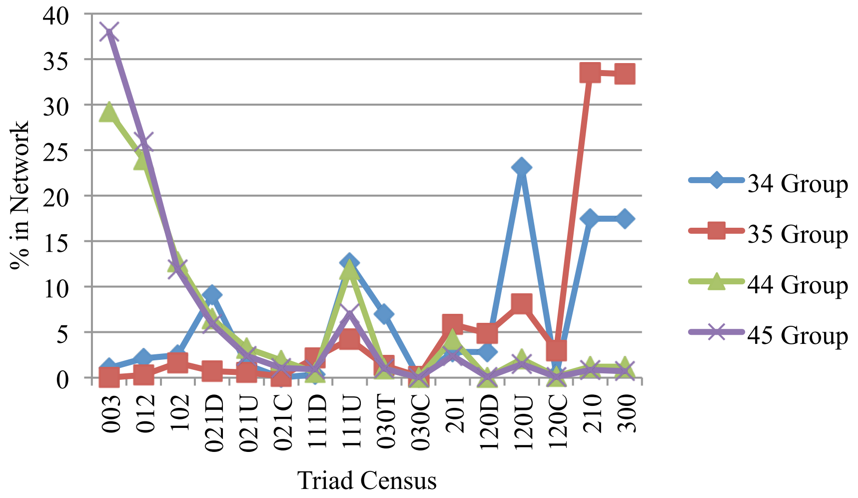

Fig.3. Triad Number Distribution for the Groups of Students

From Fig.3, Fig. 4 one can see that a triad census can be partitioned into three intervals: the first one includes three triad types {003, 012, 102}, frequency values greatly depend on the network density – significant values for a network of a low density (Δ≤ 0.5) and near zero for a dense network; the second interval includes the majority of triad types, and it is rather plain with the minimum values

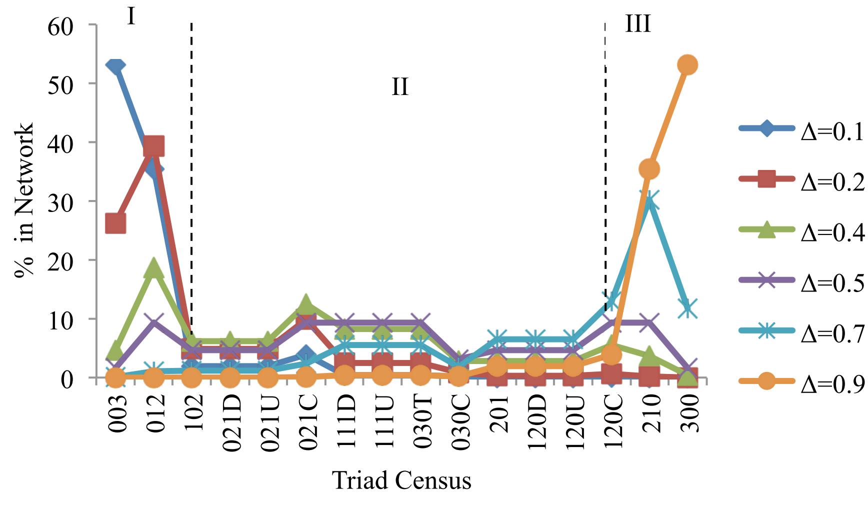

Fig.4. Triad Number Distribution for Bernouli Graphs of Various Densities

for the cyclic triads (it will be discussed further); and the third interval for three triad types {120C, 210, 300} can be considered as the mirror reflection of the first interval – significant values for a dense network and near zero for a network of a low density.

2. The main purpose of this research was considering the issue of the structural balance. The results are shown in Table 4. The 34^{th} group network is “the best” matched to the Ranked Clusters Model and the 45^{th} group is better structurally balanced in accordance with the Transitivity Model. It is worth to mention that all groups have close results against the Transitivity Model, and these scores are rather high. According to this model groups might be sustainable.

Table 4. Group Matching to the Balance Theories

№	Group of Students	Permitted Triads (%)	
		Ranked Clusters Model	Transitivity Model
1	34	64.34	83.92
2	35	50.59	84.41
3	44	56.03	81.18
4	45	61.63	88.42

3. If we compare cyclic triad number in a real group with the corresponding calculated value, we will see that cyclical structures are very seldom in the real life (their percentages are less than the corresponding theoretical ones). One can compare real and calculated values for “021C”, “030C”, “012C” triads and assess this difference.
4. It was revealed that the proportion of triads of “300” type is greater than it was calculated. This is a very special triad, as it is a clique. As Krackhardt stated, these groups (cliques) restrict people, because “group norms develop rules by which each member must play to stay a part of the group” [12]. But they are rather common in the real data.

7 Conclusions

In this research 4 small groups were under discussion. The main issue was the concept of a small group sustainability from the standpoint of the structural balance. The structural balance theories and the corresponding models form the theoretical basis of this research. It was revealed that all groups have close results against the Transitivity Model, the scores are rather high, and according to this model groups might be sustainable. If one consider the Ranked Clusters Model, one of the fourth course groups is better than the others matched to it.

During this research work the new method for social data collecting was used. In comparison with the common sociological methods of interviews, questionnaires, and observation, this method is based on information system usage. It proposes the higher degree of objectivity, but it takes more time for data collection.

It was demonstrated that a triad census is a very useful and informative characteristic, one can reveal very specific knowledge from it. Using this structural characteristic one can evaluate the structural balance in a communication network graph, and consequently predict group sustainability.

It was revealed that cyclic triads are very seldom in real communication networks, while cliques (triads with all mutual ties) are rather common.

The plans for the future research are as follows: to investigate embedded ties using triad census analysis, and to observe the first year groups in their development.

Acknowledgements: The work was supported by Act 211 Government of the Russian Federation, contract № 02.A03.21.0006

References

1. **Katz, Nancy, et al.** Network Theory And Small Groups. *Small Group Research.* Vol. 35, 3, pp. 307-332. June 2004
2. **Wellman, B.** Structural Analysis: From Method And Metaphor To Theory And Substance. In B.Wellman & S. Berkowitz (Eds.). *Social Structures: A Network Approach.* Cambridge, England: Cambridge University Press, pp. 19-61 (1988)
3. **Luhmann, N.** The Autopoiesis Of Social Systems. Eds. F. Geyer And J. Van D. Zeuwen. *Sociocybernetic Paradoxes: Observation, Control And Evolution Of Self-Steering Systems.* S.L. : London: Sage, pp. 172-192 (1986)
4. **Kovchegov, V.B.** Human Societies Group Structures Dynamics Model. *Sociology: Methodology, Methods, Mathematical Modeling.* Vol. 1, № 1, pp. 75-98 (1991)
5. **Faust, Katherine.** Comparing Social Networks: Size, Density, And Local Structure. *Metodološki Zvezki.* Vol. Vol. 3, № 2, pp. 185-216 (2006)
6. **Faust, Katherine.** Very Local Structure In Social Networks. *Sociological Methodology.* Vols. Volume 37, Issue 1, pp.209-256. December, 2007.
7. **Heider, F.** *The Psychology Of Interpersonal Relations.* New York : John Wiley, 1958.
8. **Cartwright, Dorwin And Harary, Frank.** Structural Balance: A Generalization Of Heider's Theory. *Psychological Review.* Vol. 63, № 5, pp. 277-293 (1956)
9. **Johnsen, Eugene C.** Network Macrostructure Models For The Davis-Leinhardt Set Of Empirical Sociomatrices. *Social Networks 7:203–24.* № 7, pp.203-224 (1985)
10. **Davis, James A.** Clustering And Structural Balance In Graphs. *Human Relations* . Vol. V.20, pp.181-188 (1967)
11. **Davis, , James A. And Leinhardt, Samuel.** The Structure Of Positive Interpersonal Relations In Small Groups. [Book Auth.] Morris Zeldith Jr., And Bo Anderson Edited Byj Oseph Berger. *Sociological Theories In Progress.* Boston : Houghton Mifflin, Vol. 2, pp. 218–251 (1972)
12. **Krackhardt, David.** The Ties That Torture:Simmelian Tie Analysis In Organisation. *Research In Sociology Of Organisations.* Vol. 16, pp. 183-210 (1999)
13. **Holland, P. W. And Leinhardt, S.** Dynamic-Model For Social Networks. *Journal Of Mathematical Sociology.* № 5, pp. 5-20 (1975)
14. **Shizuka, Daizaburo And Mcdonald, David B.** A Social Network Perspective On Measurements Of Dominance Hierarchies. *Animal Behaviour.* № 83, pp. 925-934 (2012).
15. **Uddin, Shahadat And Hossain, Liaquat.** Diad And Triad Census Analysis Of Crisis Communication Network. *Social Networkng.* № 2, pp. 32-41 (2013)

16. **Phellas, Constantinos N., Bloch, Alice And Seale, Clive.** Structured Methods:Interviews, Questionnaires And Observation. 3rd Revised Edition. [Book Auth.] Clive Seale. *Researching Society And Culture.* S.L.Sage Publications Ltd, pp. 181-205.
17. Total Quality Management. *Open Electronic Paper Forum.Msk.Ru.* [Online] [Cited: 12.12. 2015.] Http://Forum-Msk.Org/Material/Economic/627 694.Html.
18. **Berg, Dmitry B. And Zvereva, Olga M.** Identification Of Autopoietic Communication Patterns In Social And Economic Networks. [Book Auth.] Natalia Konstantinova, Alexander Panchenko, Dmitry I. Ignatov,Valeri G. Labunets (Eds.) Mikhail Yu. Khachay. *Analysis Of Images, Social Networks And Texts. 4th International Conference, Aist 2015 Yekaterinburg, Russia, April 9–11, 2015. Revised Selected Papers.* S.L.Springer International Publishing Switzerland, pp. 286-294 (2015)
19. **Radaev, V.V.** Market As A Web Of Social Networks. *Russian Journal Of Management.* Vol. 6, № 2, pp. 47-54 (2008)
20. **Robins, Garry.** A Tutorial On Methods For The Modeling And Analysis Of Social Network Data. *Journal Of Mathematical Psychology.* № 57, Pp. 261–274 (2013)
21. **Skvoretz, John, Fararo, Thomas J. And Agneessens, Filip.** Advances In Biased Net Theory: Definitions, Derivations, And Estimations. *Social Networks.* № 26, pp. 113-139 (2004)