RAI – Revista de Administração e Inovação
ISSN: 1809-2039
DOI:
Organização: Comitê Científico Interinstitucional
Editor Científico: Moacir de Miranda Oliveira Júnior
Avaliação: Double Blind Review pelo SEER/OJS
Revisão: Gramatical, normativa e de Formatação

ACTOR CENTRALITY IN NETWORK PROJECTS AND SCIENTIFIC PERFORMANCE: AN EXPLORATORY STUDY

André Luiz Lemes Alarcão
Mestre em administração pela Universidade Metodista de Piracicaba – UNIMEP
andre_alarcao@yahoo.com.br (Brasil)

Mário Sacomano Neto
Doutor em Engenharia de Produção pela Universidade Federal de São Carlos – UFSCar
Professor da Universidade Federal de São Carlos – UFSCar
msacomano@ufscar.br (Brasil)

ABSTRACT

This study analyzed the relationship between actor centrality of Network Projects and scientific productivity performance using a method known as Social Network Analysis (SNA). SNA and its respective properties are able to analyze actors´ positions in the structure and existing social interactions in networks. Thus, this method generates indicators to understand the format of collaborative structures of projects and their respective performances in scientific productivity. In order to carry out this proposal, models for multimodal analysis were used, taking into consideration different centrality measures. The behavior of centrality metrics has proven to be significantly different for analyses. Furthermore, the correlations between these metrics and scientific productivity performance have shown to be important in achieving project goals. This shows that the more centrality there is, the greater the chance the project has to achieve its goals.

Keywords: Social Network Analysis; Centrality; Network Projects; Scientific Productivity.

1. INTRODUCTION

Social Network Analysis (SNA) is based on methods deriving from graph theory (KILDUFF & TSAI, 2003, p. 38) and can organize structures and interactions from actors and represent them in a graph. SNA also generates individual indicators from actors or even groups and networks as a whole. These indicators can associate the nature of the structures and relations from the network to phenomena, such as power, knowledge transmission, information flow, etc. (MARTELETO, 2001, p. 72). According to Freeman (1979) SNA is a theoretical approach of a multidisciplinary nature, such as: sociology, anthropology, mathematics, statistics and computing.

According to Borgatti & Everett (1997), SNA studies attributes of pairs of individuals (or dyads), sub-groups or networks whereas in traditional social science the focus is on attributes of individuals. SNA examines structural and relational aspects in dyads, sub-groups and relationship networking (SACOMANO NETO & TRUZZI, 2009) and is also known as a meso level of analysis method. Borgatti & Everett (1997, p. 243) also highlight the importance of “pairs of individuals” in SNA, which they call *dyadic attributes*, instead of focusing on the individual itself.

As it is an approach that focuses on positioning as a technique for network studies, Borgatti *et al.* (2009, p. 901) state that the fundamental axiom of SNA lies in the concept of structures, relative to the actors´ positions. According to these researchers, the actor (node), the results and the characteristics of a network depend on this positioning (BORGATTI *et al.*, 2009, p. 902). The level in which the structure (or positioning) determines the importance of an actor (node) in a network is called centrality.

Specifically regarding collaborative environments of R&D performance, the occurrence of multiple forms of productive and technological cooperation is a recurrent theme in different approaches of Industrial Economics (BRITTO, 2002). These studies address the agglutination of skills and greater exchange of information with the R&D process (BRITTO, 2002). However, little is seen concerning how these collaborative environments influence the productivity of R&D structures (MOTE, 2005).

In this study, we attempt to find elements that enable us to clarify the dynamics of collaborative environments. In the R&D environment at Embrapa (*Empresa Brasileira de Pesquisa Agropecuária* - The Brazilian Agricultural Research Corporation), there are ways of organizing scientific research that encourage cooperative relationships to meet this demand, which are called Network Projects (NP).

This is bureaucratic because it involves normative and social formalization (CARACTERÍSTICAS, 2004), with individuals in leadership roles, characterizing the structure as interorganizational relations. In any R&D environment at Embrapa, there are actor dynamics and responsibilities for the benefit of research developed by social networks.

Each Network Project (NP) is based on a macroprogram, a management tool that conducts the operation of the company´s R&D program to obtain results that attain the technical goals. Each project consists of Research Units (RU) comprising the framework of institutions that are responsible for the activities. These activities are organized logically in a structure called Action Plan (AP) to obtain specific results expected by the project. This study specifically investigates Macroprogram 2, a portfolio that includes projects with network structures.

Embrapa has an R&D management model, according to which research projects use various actors to produce results to reach technical goals. Thus, the projects are supported by the multi-institutional and multidisciplinary approach of the actors involved. These projects generate numerous research networks with various actors, nodes and links. However, there are no systematic assessments of these networks in the company using SNA. Network measures for R&D need to be constructed so as to provide a more appropriate reading of the relationship between project structure and results. Therefore, shedding light on how network relations, specifically connection designs, have impacted the effectiveness of the company´s research results.

The main question to be addressed in the study is the following: **"Does actor centrality of Network Projects at Embrapa influence scientific productivity?"** This has implications concerning SNA measures: project structure centrality. There is no knowledge about how the centrality measures of Network Projects at Embrapa can influence the scientific productivity of the networks.

Borgatti *et al.* (2009, p. 901) highlight that the key to SNA is to understand the structural characteristics, the actors´ positions and dyadic properties. In this study, this structural term is limited to relations, focusing on the actors´ positions. As an extension of the main question, the following question arises: **"Do adjacent interactions of the actors involved in the network influence the performance of the Network Project (NP) and these actors´ scientific productivity?"**

The participants in a network may or may not have connections with other actors. When they do exist, this connectivity may be direct (also called adjacent) or even indirect.

Sometimes some actors may take on intermediary positions, exercising relative control within the universe of a whole project. Considering this, the following question arises: **"Is there a relationship between the intermediation of the actors and R&D performance in terms of these actors´ scientific productivity?"**

According to Cross & Parker (2004, p. 34), peripheral actors are those that have few connections. For these authors, this position may reflect the degree of motivation of the individual or even the little time they participate. These individuals may have a relative degree of independence in choosing (CROSS & PARKER, 2004, p. 34). This distance for the rest of the network can also denote a greater availability of suitable paths of information flow (STEPHENSON & ZELEN, 1989). Along the same line of reasoning, the more available paths there are to access other individuals, the more central this actor is. To address this issue, the following question arises: **"Does a greater availability of paths to enable access to other individuals influence these actors´ performance of scientific productivity?"**

According to Rossoni, Hocayen-da-Silva and Ferreira Jr. (2008, p. 35), the underlying assumption is that knowledge is constituted by the social environment and influenced by peers who make up an arrangement. Considering this, not only are relations observed, but also the structure which affects scientific literature. Mizruchi (2006) has the same understanding, whereby research in social networks attempts to assess the structure of the relations. Along these lines, the main objective of this study is to analyze the relationship between actor centrality of the Network Projects at Embrapa and the performance of the project in terms of scientific productivity. The propositions of the study are as follows:

- The greater the Degree Centrality (DC) of the actors involved in the projects, the greater the performance in scientific production. This hypothesis is based on the ability of actors, who have more adjacent relationships, having access to a larger number of individuals and, hence, a greater multidisciplinary structure;
- Intermediary actors perform better in scientific production projects as they ensure access to the circulation of relevant information to the network; and
- The closer the actors are, the better the scientific production project performance is, as they are more available to access other actors in the network. It is considered, therefore, that the actors who are more likely to transfer and receive information from the whole project are those who have the largest number of paths in the network.

It should be mentioned that the actors in this study are the Research Units (RU) and the Action Plans (AP) of the Network Projects.

According to Wasserman & Faust (1994), Hanneman and Riddle (2005) and Borgatti et al. (2009), there are various centrality metrics used. Three measures are recurrent in studies assessing centrality (HANNEMAN & RIDDLE, 2005) and are also addressed in this study: Degree Centrality, Betweenness Centrality and Closeness Centrality. This study took the following into consideration: Degree Centrality which is based on adjacent relationships; Betweenness Centrality which reflects the intermediation level of the structure; and Harmonic Centrality that seeks to understand the actors´ ability to be near the rest. These measures will be defined later on.

2. SOCIAL NETWORK ANALYSIS AND CENTRALITY

Studies about centralities compare the approximate central position of the actors to a relationship. Centrality indicators, using social network analysis, can investigate the degree of network connectivity, individuals with the most and least interactions, the intermediation of some actors in relationships between individuals and the closeness between the individuals who interacted (ALEJANDRO & NORMAN, 2005, p. 1). There are three metrics suggested by Freeman (1979, p 220), which are used in this study:

- Degree Centrality or DC: this is a measure that reflects the direct relational activity of an actor by measuring the number of direct connections each actor occupies in a relationship (WASSERMANN & FAUST, 1994, p. 27). According to this measure, the actor who occupies the central position is the one with the largest number of direct connections with other actors. This measure defines the degree of participation of each actor in relation to the total number of ties between the actors of the network (BORGATTI & EVERETT, 1997, p. 254). This measure indicates that a high degree of centrality reflects in the increased participation of the actor in the network. In this study, DC is considered the Degree Centrality;
- Harmonic Closeness Centrality or HC-c: To set this metric, the geodesic distance term is defined as the relationship between actors determined by the number of ties that exist in the shortest pathway between them. Closeness centrality measures how close an actor is to the other actors in the network (BORGATTI & EVERETT, 1997, p. 254). Freeman (1979) proposed this measure with the aim of measuring the ability of autonomy or independence of the actors. The higher the index, the more distant an actor is from the other actors. Thus, it follows that in this case, the distance is measured instead of the closeness. The hypothesis affirms that the more distant, the more autonomous an actor can be. To calculate the Closeness Degree, the geodesic distance of the actor in relation to all other actors in the network is added together, and then inverted, as the more distant, the less closeness (BORGATTI & EVERETT, 1997, p. 254). According to Scott (2004), this is a measure which may be indicated for global knowledge of network participants. In this study, we use the nomenclature "HC-d" to

represent the Harmonic Closeness Degree. This abbreviation 'HC' comes from *harmonic centrality* or Stephenson and Zelen´s *information centrality* (HANNEMANN, 2008), which addresses centrality as the average of close distances of the participants of the arrangement (HANNEMANN, 2008). According to Stephenson & Zelen (1989), who were the creators of this measure, Closeness Centrality considers ties as geodesic paths, while Information Centrality considers that to constitute certain information, the network can use any standard or available path (not always the shortest), and therefore use measures instead of geodesic paths.

According to these authors, this metric can be considered as another closeness measure, which deals with a "harmonic measure" of short connection paths between the actors (STEPHENSON & ZELEN, 1989). As in Tomaél´s research (2006), for this study, it is considered that the actors who are more likely to transfer and receive information from the whole project are those who have the largest number of paths in the network. The higher the index is, the higher number of paths in the network where an actor is connected to other actors.

- Betweenness Centrality or BC: The Betweenness Degree is defined as the number of geodesic distances that go through a given actor, weighed inversely by the total number of distances equivalent to the same two actors, including those that do not go through the given node (BORGATTI & EVERETT, 1997, p. 256). This is a measure aimed at measuring the intermediate positions and can be used in coordination assessments or even to control relationships. The hypothesis is that the more an actor is in intermediate positions, the more it is found in positions suitable for controlling due to the possibility of accessing information (LEMIEUX & OUIMET, 2008, p. 26).

3. PROJECT STRUCTURE AND SCIENTIFIC PRODUCTION FROM EMBRAPA

In terms of outlining the object of study, Embrapa projects were chosen where the study was developed. MP2 projects are structures on the network, entitled 'Competitiveness and Sustainability Sector', requiring structures of complex institutional projects (PRONAPA, 2007). Embrapa has 240 valid MP2 projects (PRONAPA, 2007, p. 69). To carry out this work organization, resource sharing, human skills and intra-organizational infrastructure, as well as partners are recommended (PRONAPA, 2007, p. 30).

Concerning MP2 projects, partners´ interaction and integration within or external to the Embrapa units are required settings. The MP2 Network Project has the following structural elements:

- Management Plan (MP): This aligns activities and actions to achieve the objectives;
- Action Plans (AP): Coordinated sets of efforts that transcend disciplinary boundaries and, often the technical capacity and infrastructure of a Research Unit (RU). Each AP comprises activities ordered logically, which are limited in time and are necessary to achieve results; and
- Activities (ATV): Operating determinations of research projects carried out by the participating Research Units and consecutively by leaders and participants recommended by these units. Each activity described in the Action Plans has a Research Unit, which is responsible for operations in R&D, or even managerial action in the case of the Management Plan.

To analyze the complexity of these projects, different SNA centrality measures using data obtained from the Network Projects will be used. To generate SNA measures, all the components involved in the projects and their relationships will be analysed in order to observe the closeness between these measures and scientific productivity. The following are components of the Projects: RU, MP, AP and ATV.

Concerning the collection of scientific production indicators, bibliometrics is the best known technique both nationally and internationally and entails identifying published studies. According to Macias-Chapula (1998, p. 137), this indicator "reflects the products of science, measured by counting the studies and the type of documents (books, articles, scientific publications and reports)." This study adopted scientific productivity as those productions that were submitted and accepted in national and international journals, publications in books and book chapters.

In addition to this variable, events found in technical and scientific conferences are also considered as scientific productivity indicators of Network Projects of the study, because they are considered as a quantitative category of intellectual production. Another variable used as scientific productivity was the number of products produced by the projects. Embrapa considers products as the result of different factors.

According to SIGLAS (2004), a product is all the knowledge and technology that has physical existence. Technological products such as seeds, machinery, animal breeds, as well as magazines, books, videos, CD ROMs and others are included in this format (SIGLAS, 2004).

Finally, achieving project goals is considered a variable. Network Projects aim to achieve a certain goal and, therefore, assessing the achievement level of goals relative to the different degrees of centrality is considered.

This study addresses SNA considering the analytical aspect of localized or immersed actors (nested) in all that comprises research groups.

4. METHODOLOGY

This study can be characterized as quantitative, descriptive and exploratory. In this pattern, the project centralities in question were shown in the SNA using the UCINET 6.0 software.

For the methodological development, network centrality measures were considered as independent variables, and scientific productivity measures as dependent variables.

4.1 Presentations of analysis categories

Some morphological elements of networks, such as nodes, positions, connections and flows will be used to understand the actors involved and their interactions. BRITTO (2002) defines a 'node' as the basic unit of networks. For this study, this actor is considered as the Research Units and Action Plans that make up the project. These research units are organizations which officially take part in the Action Plan of the Network Projects at Embrapa.

To designate the 'positions', the position of 'Research Activities' in the Network Projects will be adopted , e.g. in which level the research activity is included in the projects, i.e. in which Action Plan it is located.

This will depend on the structure of the project that defines the organizational logic of the Activities. These Activities are operational determinations of the projects carried out by the participating Research Units (RU) and consecutively by the leaders and participants recommended by these units.

It is intended at this level of analysis to verify the relationship structures of the networks, and for this, models capable of multi-modal analyses (2-modes) were adopted, varying according to the type of social entity involved. That is, for a certain relationship matrix, the following form of analysis was adopted:

- 2-modes: A two-mode networks, or two modes, is a network that has two distinct sets of actors, with particular attributes for each set (HANNEMANN & RIDDLE, 2005). This type of network will be used to describe matrices of relationships between the Research Units, Action Plans and the occurrence that that particular Research Unit is responsible for the activities.

4.2 Data analysis and collection

Data was collected by extracting it from the eight Network Projects available on the Embrapa Management System databases. These projects are under the heading of "genetic improvement", due to accessing consolidated project data from Macroprogram 2 and final data from 2007 to 2008. Based on the collection, the data were tabulated using Notepad version 5.1. The data were processed using UCINET and NETRAW.

The choice of the topic "genetic improvement " in the Embrapa portfolio was based on three factors: i) because it is a topic in the context of Network Projects, an obligatory requirement of Macroprogram 2; ii) because it is a very important topic for the company's mission; and iii) because it is one of the topics which has been on the networks for a long time.

To use bimodal matrices, the projects were divided up to determine the extent of the relationships. In other words, they addressed the existing relationships in the projects. Data were analyzed from the statistics generated by SNA.

5. RESEARCH RESULTS

Descriptive and inferential statistical studies were conducted to verify if there were any significant differences between the Degree Centrality (DC), Betweenness Centrality (BC) and Harmonic Closeness Centrality (HC-c) metrics.

By the descriptive statistics (Table 1), it can be observed that the DC, BC and HC-c metrics possibly follow different distributions, as the values of the core measures (mean and median) are relatively different. The BC metric showed much lower values for the mean and quartiles, while the DC metric presented intermediate values, and the HC-c showed higher values. The standard deviation observed indicates that the metrics studied have similar variability.

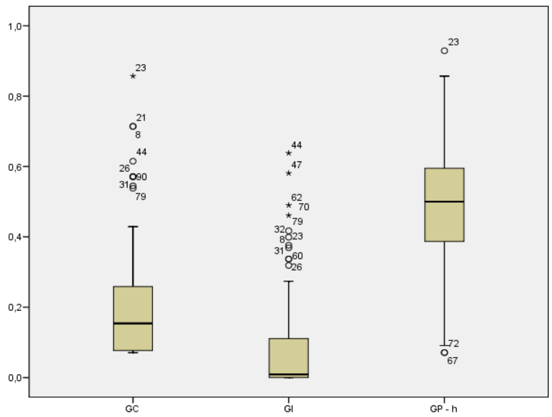
Table 1 Descriptive measures for the DC, BC and HC-c variables

		DC	BC	HC-c
N	Valid	101	99	101
	Absent	0	2	0
Mean		0.20597	0.08238	0.49271
Median		0.15400	0.00900	0.50000
Mode		0.071	0.000	0.607
Standard Deviation		0.167574	0.140324	0.167545
Minimum		0.071	0.000	0.071
Maximum		0.857	0.638	0.929
Percentile	25	0.07700	0.00000	0.38700
	75	0.25850	0.11500	0.59500

Source: Elaborated by authors

This descriptive analysis was studied in more depth by using a box-plot. According to Moore (2005, p. 35), these graphs enable us to observe a central box bounded by Q1 and Q3 quartiles, a straight line paired with the median, and straight lines for observations of higher or lower values. According to Figure 1, it can be observed that the central rectangles of these diagrams, which account for 50% of the central distribution values present a visual differentiation between the three metrics. It is also worth mentioning the significant occurrence of the data called *outliers*, which clash between the metrics. This is data that clash with others (Moore, 2005) and which occur significantly in the DC and BC.

Figure 1 Box-plot of the DC, BC and HC-c variables



Source: Elaborated by authors

The variables studied appear to belong to different distributions, and have HC-c values higher than the others, and the BC variable is more concentrated at lower values. The BC has a very flat graph between the minimum and the median, which shows that there is too much concentration in a small space of values. Thus, the average BC would be even lower if it did not have so many outliers, demonstrating how the average distanced itself from the median for this metric. The outliers eventually increased the median; in this case, the median was the best measure of central tendency.

However, the HC-c has a more symmetrical distribution whose mean is relatively close to the median. Quartiles of these metric values do not present such close values as the BC, in which the first quartile tends toward the mean.

Finally, the DC metric is relatively symmetrical, and the mean and median are not very distant. According to the figure, the 25% lower DC values are very close to the DC. As for the BC, this metric has many outliers and this contributes to increasing the average.

5.1 Evidence that the centralities do not come from the same distribution

For confirmation about the differences between these metrics, we chose to use the Friedman Test (Table 2). It is a non-parametric test that uses multiple comparisons of variances (MOORE, 2005, p. 540). As the measures generated by DC, BC and HC-c are for the same actors, it was considered that in this test data are paired or dependent.

Table 2 Friedman Test showing differences between DC, BC and HC-c

Friedman Test	
N	99
Valor-p	0.000

Source: Elaborated by authors

According to the results obtained from the Friedman test, it was found that there is a significant difference between the variables studied, confirming what was observed by the descriptive statistics. It can be observed that the p-value of the Friedman test was less than 0.001 confirming that the three metrics have different distributions. The test considered that the three samples from the population with the same distribution as H0 and H1 was the opposite. This test considered that the metrics do not belong to the same population, i.e. they are different.

5.2 Analysis of pairing

Data distribution was based on three related samples (DC, BC and HC-c) in the eight conditions tested. After proving there was a difference of distribution between the metrics, we tried to show where the difference was. The question was whether the three were different, or only one of them. Owing to the fact that in SPSS (Statistical Package for the Social Sciences), there is no multiple comparison for paired tests, the alternative was to compare 2-2 with paired tests. Therefore, two tests were run to see where the difference was: the Wilcoxon test and the Paired t-test (Table 3).

Table 3 Paired t-test (comparison 2-2)

Paired Differences								
Comparison Index								
	Mean	Standard Deviation	Standard Error	Lower L	Higher L	t	df	p -value
1 DC - BC	0.124859	0.11833	0.011893	0.101258	0.148460	10.499	98	0.000
2 DC - HC-c	-0.286743	0.108744	0.010820	-0.308210	-0.265275	-26.500	100	0.000
3 BC - HC-c	-0.410182	0.143392	0.014411	-0.438781	-0.381583	-28.462	98	0.000

Source: Elaborated by authors

This test is used to compare the paired or dependent groups regarding some quantitative variable (MOORE, 2005, p. 553). The p-value was lower (<0.001) than the significance level of 0.05 for all the comparisons 2-2, showing that there is a difference among the groups.

After this stage of analysis, a normality test was used to compare the frequency curves. According to Cirillo and Ferreira (2003), identifying normality in data is generally done using graphs. Simply observing the graphs is not sufficient, especially in the multivariate case, and specifically in situations of many variables.

5.3 Tests for independent samples

According to the Kolmogorov-Smirnov Normality Test, it was observed that only the HC-c can be from a normal distribution, as it was the only one in which the null hypothesis of normality was not rejected. Thus, the Pearson correlation test cannot be used. The next step was then to carry out a non-parametric correlation test, the Spearman Correlation test (Table 4).

Table 4 Spearman Correlation test

		DC	BC	HC-c
DC	Correlation Coefficient	1.000	0.826	0.803
	p-value	-	0.000	0.000
	N	101	99	101
BC	Correlation Coefficient	0.826	1.000	0.754
	p-value	0.000	-	0.000
	N	99	99	99
HC-c	Correlation Coefficient	0.803	0.754	1.000
	p-value	0.000	0.000	-
	N	101	99	101

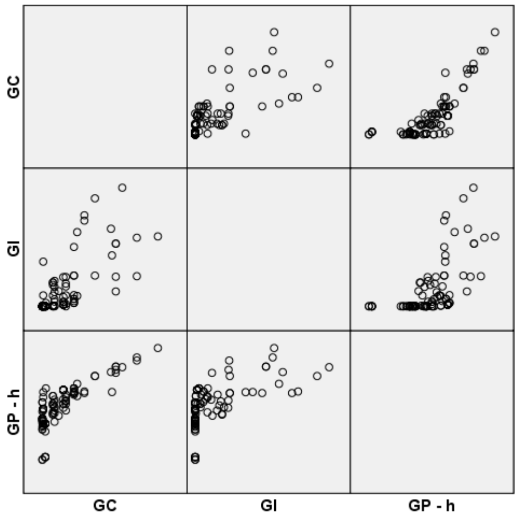
Source: Elaborated by authors

Based on the correlation tests, it can be observed that the metrics have a positive correlation (correlation coefficient> 0.7), and significant, as the p-value of the tests was less than the significance level (0.05). Therefore, although the tests above show that the three metrics have different distributions, it can be observed that with the previous test they are correlated. That is, when an individual has a high value in the BC metric, this will more likely have a high value in the other metrics.

Referring to Figure 2, the behavior of the variables together can be observed using the statistical tool: *scatter plot*. It can be seen that the tables which show points forming an increasing straight line are from the metrics that have a positive correlation, i.e., when one increases, the other also tends to rise. However, if the line decreases, then there is a negative correlation indication, i.e.

when one decreases, the other increases and vice versa. It was observed in Figure 2 that the dispersion approaches a straight line, and the closer it is, the more linear the relationship is.

Figure 2 Dispersion Matrix for the DC, BC and HC-c variables



Source: Elaborated by authors

In quadrants (DC x DC, DC x DC, BC x-CC and HC-c c x BC) of the Dispersion Matrix, the data that are positioned in a dispersed way show disagreement between the distributions. It can be observed that there was more linearity in the patterns of the graphs of HC-c metrics compared with the graphs of the BD and DC metrics. For the quadrants without the HC-c metric, the graphs did not show a more stable form.

5.4 Statistics between metrics and productivity projects

The aim of observing the matrix indexes was to verify a possible correlation between the metrics and productivity of the projects. Based on this step, we attempted to use the descriptive and inferential statistics to check for correlations with the performance of the projects.

Thus, evaluating the correlation between the centrality measures and project performance took into account the data extracted from the Final Reports of the Network Projects. The variables considered were:

- Goal: According to the leader, success was achieved according to the initial planning of the project;
- Total number of results: Reports of research results;

- Publications: Number of publications submitted and approved in technical and scientific journals; and
- Events: Number of events carried out during the project.

Table 5 Production and achievement of project goals

Abbreviations	Goals (%)	Total number of results	Publications	Events
A	81.25	14	12	50
B	98.00	25	20	22
C	91.00	11	61	29
D	75.00	1	5	9
E	65.00	11	14	37
F	77.00	12	38	32
G	70.00	8	36	31
H	63.25	15	33	69

Source: Elaborated by authors

Concerning Table 5, the data presented represent the scientific performance extracted from Final Reports of the eight projects studied. Each project is represented by a letter, the percentage of achieving the targets, the total number of products, publications and events.

As shown in Table 6, it was observed that the HC-c metric obtained a higher correlation coefficient with the Goal to 0.536 and was the only one that showed a significant correlation (p-value <0.001). It can also be observed that the Network Project with higher achievement goals have higher HC-c. For other productivity variables for HC-c, there was a small positive correlation with the Total Results (0.010) and a small negative correlation with the variables Publications and Events. The negative correlation coefficient demonstrates that the higher a variable, the smaller the other one will be. In both results, "Total Number of Results" and "Publications and Events", the indexes were considered inconclusive

Concerning the other BC and DC metrics in Table 6, all the correlations observed obtained values below 0.243 (DC x Goals) and above -0.064 (BC x Publications) for negative indexes, which does not show that there are correlations between the DC and BC metrics with productivity and achieving goals.

Table 6 Correlation between the metrics and productivity achieving goals measures

Metrics	Goals	Total Number of Results	Publications	Events
DC	0.243 p-value 0014	0.072 p-value 0.472	-0.079 p-value 0.432	-0.075 p-value 0.454
BC	0.174 p-value 0.086	-0.081 p-value 0.423	-0.064 p-value 0.530	0.085 p-value 0.401
HC-c	**0.536** **p-value <0.001**	0.110 p-value 0.272	-0.221 p-value 0.026	-0.305 p-value 0.002

Source: Elaborated by authors

The first analysis only considers the only analysis units (Research Units and Action Plans). In an attempt to compare information, afterwards the analyses consider only Research Units as analysis units, excluding the Action Plan actors, as shown in Table 7.

Table 7 Correlation considering only the Research Units as an actor of interest

Metrics	Goals	Total Number of Results	Publications	Events
DC	0.178 p-value 0.190	1.124 p-value 0.361	0.086 p-value 0.528	0.051 p-value 0.711
BC	0.096 p-value 0.491	0.022 p-value 0.875	0.174 p-value 0.207	0.125 p-value 0.368
HC-c	**0.479** **p-value <0.001**	0.283 p-value 0.035	-0.110 p-value 0.420	-0.104 p-value 0.443

Source: Elaborated by authors

These analyses corroborate previous analyses. Once more there were low correlations for the variables of interest, because it was the Goal variable that HC-c showed the highest correlation coefficient to 0.479. All the other correlation indicators were discarded as evidence of correlation.

The second analysis only considered the Action Plans (AP) as an actor, and as shown in Table 8, once again the correlations were not significant for the variables of interest. The Goal variable had the highest correlation with HC-c and had a correlation coefficient of 0.563, followed by a negative correlation from Events, -0.484. Both correlations presented a p-value <0.001 proving that there were significant correlations.

As the aim is to compare the various designs of network projects, taking the centrality of each element of the project as a measure for these designs in relation to this, we proposed to use a single measure for each project.

Table 8 Correlation only considering Action Plans as actors of interest

Metrics	Goals (Effective Measuring Tool for the Project)	Tital Number of Results	Publications	Events
DC	0.416 p-value 0.005	0.013 p-value 0.932	-0.357 p-value 0.016	-0.307 p-value 0.040
BC	0.349 p-value 0.019	-0.148 p-value 0.330	-0.346 p-value 0.020	-0.315 p-value 0.035
HC-c	**0,563** **p-value <0.001**	-0.185 p-value 0.223	-0.249 p-value 0.099	**-0.484** **p-value <0.001**

Source: Elaborated by authors

The proposal to use a single measure considered the measures synthesized using the means and afterwards, the standard deviation. For each project, the means and the standard deviation of DC, BC and HC-c metrics were extracted.

5.5 Centrality means of each project

This choice had the assumption that a project structure with higher means would perform better. Based on this mean, a correlation test was used between it and the variables of interest. In Table 9, it can be observed that there was a correlation value of 0.833 in HC-c with the Goal variables, which can be considered a strong positive correlation. It should be mentioned that no p-value showed a significant correlation (more than 0.010).

Table 9 Correlation of productivity variables and centrality means

Metrics	Goals	Total Number of Results	Publications	Events
DC	0.595 p-value 0.120	0.108 p-value 0.799	-0.095 p-value 0.823	-0.119 p-value 0.779
BC	0.333 p-value 0.420	-0.587 p-value 0.126	-0.119 p-value 0.779	-0.357 p-value 0.385
HC-c	0.833 p-value 0.010	-0.132 p-value 0.756	-0.262 p-value 0.531	-0.524 p-value 0.183

Source: Elaborated by authors

The same analysis was made separating the actor as Research Unit, and as Action Plan respectively presented in Tables 10 and 11.

Table 10 Correlation of productivity variables and centrality means only considering the Research Units

Metrics	Goals	Total Number of Results	Publications	Events
DC	0.190 p-value 0.651	0.108 p-value 0.799	0.500 p-value 0.207	0.214 p-value 0.610
BC	-0.286 p-value 0.493	-0.263 p-value 0.528	0.786 p-value 0.021	0.286 p-value 0.493
HC-c	0.810 p-value 0.015	0.156 p-value 0.713	-0.143 p-value 0.736	-0.452 p-value 0.206

Source: Elaborated by authors

Table 11 – Correlation of productivity variables and centrality means only considering the Action Plans

Metrics	Goals	Total Number of Results	Publications	Events
DC	0.738 p-value 0.037	0.108 p-value 0.799	-0.500 p-value 0.207	-0.524 p-value 0.183
BC	0.416 p-value 0.233	-0.072 p-value 0.866	-0.714 p-value 0.047	-0.548 p-value 0.160

HC-c	0.810 p-value 0.015	0.168 p-value 0.691	-0.333 p-value 0.420	-0.476 p-value 0.233

Source: Elaborated by authors

When separated, the Actor and Plan groups showed the same behavior, and there was a positive correlation of 0.810 with the variable "Goals", the most significant correlation value, however it was not significant (p-value 0.015).

6. RESULTS ANALYSIS

As part of the method that make up these structural characteristics (WASSERMANN & FAUST, 1994), the centralities in this research have analytical differences in their use, which infers that the complexity of the project structures can influence the definition of which centrality is to be used in different analytical contexts in networks. Taking this into account, it was shown that the Degree Centrality (DC), Betweenness Centrality (BC) and Harmonic Closeness Centrality (HC-c) metrics follow different distributions between them, and are therefore different for analyses. The Betweenness Centrality (BC) and Degree Centrality (DC) are not shown to be stable compared with HC-c due to the great variability of the data. In this respect, there is a significant occurrence of outliers in DC and BC, which proves that the most appropriate metric for Network Project studies from Embrapa was the Harmonic Closeness Centrality (HC-c).

Another finding is that there is a correlation between the three metrics studied. This means that the higher the DC, the higher the BC and HC-c, but when compared individually there is more linearity in the patterns of the HC-c metric, if compared with the BC and DC metrics. Despite this correlation and the fact that they are different metrics, it is considered that for the structures of the projects analyzed, where interactions occurred between the Research Units and Action Plans, there is a need for analytical differentiation. This fact shows how centrality is an important position in the context of networks, as highlighted by Miziruch (2007).

When dealing with actors that represent work divisions, where the Research Units comprise the Action Plans, it was shown that the three metrics authentically expressed degrees of positions needed to confirm the interactions. This reinforces that even with a more statistically balanced metric for analyses, both can show the project design and their respective interactions.

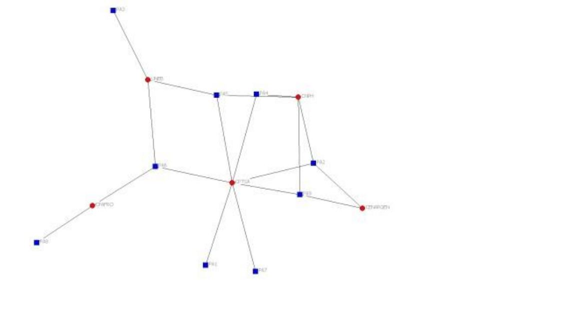
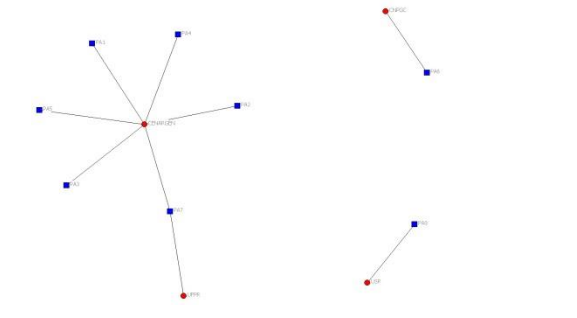
These findings presented above assessed the metric itself. To meet the objectives of the study, the project performance needed to be associated with SNA. Network projects that reach the goals set have a higher HC-c. This shows how close the actors are to the others, showing how easy it is to interact and, consequently, achieve goals. This means that individuals who have more HC-c are those that have a greater number of paths in the network, according to Stephenson & Zelen (1989), a greater chance of receiving information from the whole network.

The impacts of BC and HC-c concerning productivity (goals) may indicate that the strategy to connect APs to RUs which, in turn are well connected to other APs (> HC-c) has a clear advantage over the strategy of having Action Plans acting as "bridges" between different RU.

According to Zelen & Stephenson (1989), HC-c analyzes the information flow (TOMAEL, 2006). The analyzed networks have shown that the combination of paths between the actors may be more useful than properly checking the betweenness position (BC) or even the position that expresses adjacent relationships (DC). According to the Harmonic Closeness Degree, the actors who have higher chances of information flow are those with the greater number of paths in the network.

To illustrate this finding, the project structures can be observed, described in Figure 3.

FIGURE 3: Network Projects G and H

Project G	**Project H**
	

Source: Elaborated by authors

According to Figure 3, it can be observed that Project G has more information channels and a greater HC-c compared to Project H, that determines which actors from this structure (Project G) are more likely to receive and transmit information and do it consecutively. This more central network (HC-c) was correlated with higher levels of achievement goals.

It can be observed in Project G that there is a wider range of relationships between all the actors, which even having few Research Units (five) and two Action Plans with only one Research Unit, indicating low mobilization of actors, is a project with a higher Degree Centrality compared with Project H. According to the results of this study, it was proved that the project with a higher HC-c and consequently more information and contact flow made it easier to obtain the desired goals compared with smaller Harmonic Closeness Degree networks.

When the three metrics are compared, it is highlighted that even adopting a bimodal analysis model with actors represented by Action Plans and Research Units, it can be concluded that Freeman´s metrics (1979) of DC and BC for the work division analyses are different for analyses whereby there are work divisions with non-adjacent relationships between the Research Units.

Having the evidence that the Harmonic Closeness Degree is a feasible metric for Network Project analysis, it is understood that the availability of actors to choose from a greater number of paths to follow influences the results of the projects, or at least helps to achieve the goals set.

Regarding the Degree Centrality (DC) and Betweenness Centrality (BC), conclusive correlations were not observed with the productive performance of the projects. This finding goes against the hypothesis of the study, whereby the greater the ability of adjacent relationships, the better the benefit of the actor, and the more betweenness, the more access there is to information. It is assumed in this case that a greater number of direct actors´ relationships, showing increased DC, and greater betweenness capacity may not be conclusive because of the interoperability variable, i.e. the actors´ ability of flowing in the network between clicks (or subgroups) and the ability of actors to take advantage of the available paths. In other words, it is considered that the organizational mobility considering the availability of paths is a key factor in achieving results of Network Projects at Embrapa.

7. CONCLUSIONS

Interaction between the participants of the Network Projects at Embrapa gives rise to a structure of relationships with different positions of the actors involved. The diversity of these positions determines the multiple forms of scientific and technical performance of the projects. All the multi-institutional involvement comes from a process in which the association of skills to meet institutional goals attempts to position suitable structures to format the projects. The relationship paths are noticeable, especially in more central branches in the network universe. These paths consolidate the social structure of the projects and maintain alliances.

Not all the projects´ participating actors are socially involved in the different spheres of the research proposed by Action Plans. In other words, the scope of relations of formatting the projects does not always include all the Research Units that make up a project. Thus, there are many degrees of centralities of those involved.

The RUs that have greater centrality are those which are most involved in the research activities and are responsible for forming ties with other RUs, which are consequently well connected to other APs. That is, the most involved RUs in the Action Plans are those that have greater centrality levels, with more possibilities of institutional mobility. These actors are those from the Social Network Analysis (SNA) identified as individuals with higher Harmonic Closeness Degrees, and as the most articulate actors. These RUs are mainly responsible for “moving” information in the Network Projects.

The actors that have the greatest number of available channels for information flow receive this information mostly from the network and are more effective in achieving the goals proposed by the projects. The importance highlighted in this study for them justifies the joint effort of the project that prioritizes consolidating relationships, mostly from the participating Research Units in most of the projects´ Action Plans.

The centrality of the actors gives the project productive capacity and importance. More centrality of the actors influences sharing information, providing adequate dissemination, cooperation and establishing channels for knowledge.

It is worth mentioning that if we take into account that actor centrality in a project has a certain amount of influence on the performance, all the actors are relatively important, including those who are in more remote positions.

This centrality is a key element in understanding the collaborative processes, but it is considered to be a part of comprehending the dynamics of inter-organizational relationships, because other factors, such as cohesion should be taken into consideration.

It is important to point out that in this study it was not possible to find correlations of the metrics with the productivity of projects, concerning the Results (products), Publications and Events. It can be concluded as a limitation of the research that the amount of data was insufficient to prove this correlation or not. Further testing needs to be done using a larger number of assessed projects and more data from final report projects, statistical evidence pointed out in this study.

This limitation leads to new research. In order to continue the current research, the following should be considered: 1) It is recommended to extrapolate these analyses for more complex projects in their relationship structure (more ties and actors); 2) To assess relational variables such as confidence or perceived value among the actors of the network.

In addition to these theoretical developments, it is believed that this study is an initial instrument at Embrapa to improve the construction of Network projects in the articulation and project development stages, focusing on identifying partners, stakeholders and paths in a possible suitable structure of institutional and interpersonal relationships.

REFERENCES

ALEJANDRO, V. Á. O.; NORMAN, A. G. (2009). **Manual introductorio al análisis de redes sociales: medidas de centralidad.** Available at: <http://revista-redes.rediris.es/webredes/talleres/Manual_ARS.pdf> Accessed on 10th February, 2009.

BALESTRIN, Alsones; VARGAS, Lília Maria. (2005). A dimensão estratégica das redes horizontais de PMEs: teorização e evidências. **Revista de Administração Contemporânea,** Edição Especial, pg. 203-227.

BORGATTI, Stephen P., EVERETT, Martin G. (1997). Network analysis of 2-mode data. **Social Networks,** v.19, p.243-269.

BORGATTI, Stephen P. (2002). **NetDraw**: Graph Visualization Software. Harvard: Analytic Technologies. Available at http://www.analytictech.com/Netdraw/.

BORGATTI, Stephen P.; EVERETT, Martin G. (2006). A graph-theoretic perspective on centrality. **Social Networks,** v. 28, p. 466-484.

BORGATTI, Stephen P**.** (2009). **2-Mode Concepts in Social Network Analysis**. Available at: http://www.steveborgatti.com/papers/2modeconcepts.pdf. Accessed on 10th February, 2009

BRITTO, J. (2002). Cooperação interindustrial e redes de empresas. In: KUPFER, D. **Economia industrial**: fundamentos teóricos e práticos no Brasil. Rio de Janeiro: Editor Campus.

BURT, R. S. (1992). The Social Structure of Competition. In: NOHRIA, N.; ECCLES, R. G. (Eds.). **Networks and organizations**: structure, form, and action. Boston: Harvard Business School Press.

CARACTERÍSTICAS e Gestão dos Macroprogramas. (2004). In: **Manual do Sistema Embrapa de Gestão**. Brasília, DF.

CIRILLO, M. A.; FERREIRA, D. F. (2003). Extensão do teste para normalidade univariado baseado no coeficiente de correlação quantil-quantil para o caso multivariado. **Revista de Matemática e Estatística,** São Paulo, v. 21, n. 3, p. 57-75.

CROSS, Rob; PARKER, Andrew. (2004). **The Hidden Power of Social Networks:** Understanding How Work Really Gets Done in Organizations. Boston: Harvard Business School Press.

EMIRBAYER, M; GOODWIN, J. (1994). Network analysis, culture and the problem of agency. **American Journal of Sociology**, v. 99, n. 6, p. 1411-54.

FREEMAN, Linton C. (1979) Centrality in social networks: conceptual clarification. **Social Networks,** v.1, n.2, p.215-239.

GRANOVETTER, M. (2007). Ação econômica e estrutura social: o problema da imersão. **RAE Eletrônica**, v.6, n.1.

GRANOVETTER, M. S. (1983). The strength of weak ties: a network theory revisited. **Sociological Theory**, v.1, p.201-233.

HANNEMAN, Robert A.; RIDDLE, Mark. (2005). **Introduction to Social Network Methods. Riverside**: University of California. Available at <http:/faculty.ucr.edu/hanneman/nexttext/index.html> Accessed on 10th September, 2008.

KILDUFF, Martin e TSAI, Wenpin. (2003). **Social networks and organizations**. London: Sage.

LEMIEUX, Vicent; OUIMET, Mathieu**.** (2008). **Análise estrutura das redes sociais**. Lisboa: Instituto Piaget.

MACIAS-CHAPULA, César A. (1998). O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciencia da Informação** [online]. v. 27, n. 2, p. 134 –140. Available at <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651998000200005&lng=en&nrm=iso> Accessed on 27th March, 2009.

MARTELETO, Regina M. (2001). Análise de redes sociais – aplicação nos estudos de transferência da informação. **Ciência da Informação,** Brasília, v. 30, n. 1, p. 71-81, Jan./Apr.

MARTES, A. C. B. *et al*. (2007). Fórum: Sociologia econômica. **Revista de Administração de Empresas**, v.47, n.2, p.10-14.

MIZRUCHI, M. S. (2006). Análise de redes sociais: avanços recentes e controvérsias atuais. **Revista de Administração de Empresas**, v. 46, n. 3, p.10-15.

MOORE, David S. (2005). **A estatística básica e sua prática**. 3. ed. Rio de Janeiro: LCT.

MOTE, J. E. (2005). R&D ecology: Using 2-mode network analysis to explore complexity in R&D environments. **J. Eng. Technol. Management**, v. 22, p. 93-111.

NOHRIA, Nitin. (1992). Is a network perspective a useful way of studying organizations? In: **Networks and organizations: structure, form, and action.** Boston: Harvard Business School Press.

POWELL, Walter W.; KPOUT, K.; SMITH-DOERR, Laurel. (1994). Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. **Administrative Science Quarterly**, v. 41, n. 1, p.116-145.

PRONAPA. (2007). **Programa Nacional de Pesquisa e Desenvolvimento da Agropecuária**. Brasília: Embrapa.

ROSSONI, L.; HOCAYEN-DA-SILVA, A. J.; FERREIRA JÚNIOR, I. (2008). Estrutura de relacionamento entre instituições de pesquisa do campo de Ciência e Tecnologia no Brasil. **RAE. Revista de Administração de Empresas**, v. 48, p. 34-48.

SACOMANO NETO, M.; TRUZZI, Oswaldo M. S. (2002). Perspectivas contemporâneas em análise organizacional. **Gestão e Produção,** v.9, n.1, abr.

SACOMANO NETO, M.; TRUZZI, Oswaldo Mário Serra. (2009). Posicionamento Estrutural e Relacional em Redes de Empresas: uma análise do consórcio modular da indústria automobilística. **Gestão & Produção**, v. 16, p. 132.

SCOTT, J. (2004). **Social network analysis**: a handbook. Available at: <www.analytictech.com/mb119/tableof.htm>. Accessed on 8th Dec.

SIGLAS, (2004). Definições e Terminologias do Sistema Embrapa de Gestão – SEG. In: **Manual do Sistema Embrapa de Gestão**. Brasília, DF.

SIMMEL, G. The Triad. In: WOLFF, K. H. (Ed.). (1950). **The Sociology of Georg Simmel**. New York: Free Press.

STEPHENSON, Karen; ZELEN, Marvin. (1989). Rethinking centrality: methods and examples. **Social Networks**, v.11, n.1, p.1-37, Mar.

TOMAÉL, Maria Inês. (2006). Redes sociais: posições dos atores no fluxo da informação. **Revista Eletrônica Biblioteconomia e Ciência da Informação**, Florianópolis, 1º sem.

WASSERMAN, Stanley; KOEHLEY, Laura. (1994). Classification of actors in a social network base on stochastic centrality and prestige. **Connections: Bulletin of The International Network for Social Network Analysis**, p. 35-44.

The date of receipt: 07/29/2015

The date of acceptance: 03/31/2016