

Formal Information Flows Among Top Authorities of the Brazilian Federal Government based on Co-word Analysis of Data Published in the Official Gazette

Rafael Henrique Santos Soares ¹, Jorge H. C. Fernandes¹ Ricardo Sampaio¹

¹Grupo de Pesquisas sobre Inteligência Cooperativa em Redes Sociais Complexas – Programa de Pós-Graduação em Ciência da Informação – Faculdade de Ciência da Informação - Universidade de Brasília (UnB)
Campus Universitário Darcy Ribeiro, CEP: 70.910-900, Brasília- DF, Brazil
rafaelhss@gmail.com, {[jhcf|rbsam](mailto:jhcf|rbsam@unb.br)}@unb.br

Abstract. *This paper describes a methodology for modeling social networks of formal communications among top authorities of the Brazilian Federal Government grounded on data available in the government official gazette (diário oficial da união). The text of a large number of official publications such as presidential decrees, ministerial orders and authority nominations was analyzed for identification of citations to organizations and persons. The co-occurrence of names of persons in such publications created a network of relations among such persons. An ego-network was built around the president Dilma Rousseff. Metrics of social network analysis were collected and analyzed in an exploratory fashion.*

1. Introduction

Information flows every day in public and private agencies, either between persons either between organizations. Communication is a natural and necessary process that supports collaboration and coordination of work. Public organizations must behave in accordance with the public interest and during the last decade Brazil's federal government has strongly evolved in publishing data about high level government decisions using portals over the web and thus enhancing its transparency.

Under a political point of view, information communication defines and is defined by the interaction among personal agents in fluid networks that exhibit heterogeneous and emergent properties. People talk through a lot means, trying to get necessary information to accomplish something or to convince others that something must be accomplished.

Federal agencies at the top of the formal hierarchy, such as ministries, are public institutions whose agents naturally develop heavy political bias, but they also must act in order to implement and evaluate public policies. This way, understanding the information network formed in a political scenario of a public organization is important to understand the dynamics of power and influence of public agents in the development of such policies.

This paper postulates that by focusing on the networks established by the persons that are top government authorities makes it possible to infer on the

collaborations and opportunities that are explored by these persons, based on their structural position on the network.

However, it is hard to map this type of network given that data about communication between top authorities is often not readily available and interviewing such persons on political questions is hard and expensive to do, even with a moderated number of respondents.

This paper explores the use of data available on web sites to map networks of communications among public agents in the Brazilian Federal Administration. We collect text from D.O.U. (“Diário Oficial da União” – Federal Official Gazette) to build social networks of public agents and. Through the analysis of this network of formal communications we aim to understand how these agents interact. With the data collected we have several options of analysis, but in this case we focused on networks established around the president of republic, and in how this interaction influences other opportunities to mediate information. We also analyze the dynamics of information flow and the effect of a structural position in nominations on the government.

2. Information communication between public agents and organizations and its effects in official media

Acts published at official media such as the D.O.U. (“Diário Oficial da União” – Federal Official Gazette) are manifestation of communication between public agents and organizations. In that sense, we explore Le Coadic’s (2004, p.35) concepts of formal and informal communications, Costa’s (2013) idea of a gradual transition from informal to formal and the Spiral of Communication Model proposed by Dance (1967, p.374). Based on these concepts, a model of information communication between public organizations and agents and its manifestation in official media like D.O.U. is presented.

When discussing information communication, Le Coadic (2004, p.35) argues that the communication process can be divided into two aspects: a written/formal and an oral/informal. The author asserts that the formal process produces stable information while the informal one delivers information more subject to changes and repetition. The author also states that the informal processes precede the formal ones.

An analogy to Le Coadic’s (2004, p.35) theory can be made to understand the information communication process between public agents and organizations which also has a formal and an informal component. Decisions are made internally after conjectural analysis, discussions, studies among other supplementary processes that, depending on the matter, include public agents outside the organization. Some published decisions about collaboration deals, working groups creation or project implementations are made after months or years of negotiation and informational exchange between the public organizations involved.

Table 1 shows an adapted version of Le Coadic’s (204, p.37) scheme comparing elements of formal and informal information communication. In this table, we included two columns presenting examples of each element stated by the author in the context of the information communication between public agents that can be found in official publications like D.O.U.

Table 1 – Examples of elements found in D.O.U.

	Formal	Informal
Cited by Le Coadic	Public (big potential audience)	Private (restricted audience)
Example in D.O.U.	Published online	Phone calls, e-mails, meetings.
Cited by Le Coadic	Information stored permanently	Information not stored and not recoverable
Example in D.O.U.	Files and online repository with open access.	Sometimes some poor registries can be found, like minutes and reports, but are often well structured.
Cited by Le Coadic	Information relatively old	Fresh new information.
Example in D.O.U.	A publication at D.O.U. is a communication to society of something already done	The decision ins been built during the communication process
Cited by Le Coadic	Information proved truth. Can be trusted.	Not proved information.
Example in D.O.U.	Publicity is a requirement for an act to take effect. Therefore, it is truth.	When communicating informally, people often do not worry about accuracy.
Cited by Le Coadic	Uniform spread.	The flow of the information is chosen by the
Example in D.O.U.	Access to D.O.U is open to anyone.	Only those invited to interact have access to the
Cited by Le Coadic	Low redundancy.	Redundancy is often an important element.
Example in D.O.U.	One single publication is enough for an act to take effect.	Complex matter are discussed by various agents inside organizations with gradual involvement of upper hierarchy, requiring redundancy.
Cited by Le Coadic	No direct interaction.	Direct interaction by default.
Example in D.O.U.	Publications in D.O.U are one-way	Public agents interact while discussing.

Source: Adaptation of Le Coadic’s (2004, p.36).

Costa (2013) made a graphic view of the organizational communication channels’ classification. Figure 01 synthetizes this scenario and allows us to identify a hierarchical relation between types of channels.

Analyzing figure 01 we can see that there is no polarity between formal and informal, but, a gradient that begins at one side completely formal and ends in the other, completely informal. In the middle lays the communication channels distributed gradually in this ‘rank’ according to its ‘formality level’.

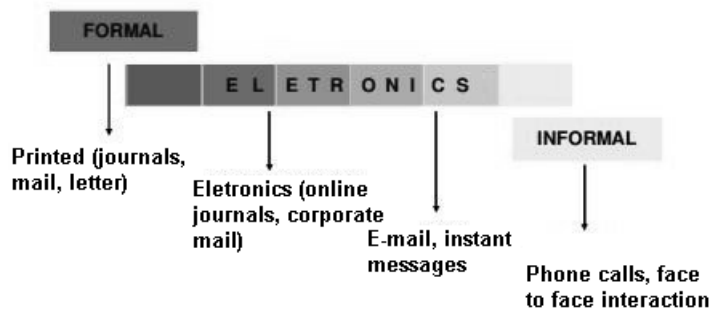


Figure 01 – Classification of communication channels.
Source: Costa (2013), translated

Dance (1967, p.378) argues that communication must be understood as a continuous and incremental process. He criticizes the linear models (for not represent feedback) and the circular ones (for not represent the effect of information exchange in individuals), and suggests a spiral model. The author believes that we should not use lines or circles but a helix (or a spiral) instead as the basic shape for represent human communication. The spiral, he says, is better for it shows the direction of the

communication flow, like the line, the feedback, like the circle, and the incremental impact that communication has on itself, making it bigger and aggregating more information every iteration.

McQuail and Windahl (1993) emphasize the dynamic aspect of Dance's model comparing to the "frozen picture" of the process provided by traditional communication models. The authors point out, however, that Dance (1967) did not provide a detailed tool or model for deep analysis but, instead, identified a core characteristic of the communication process that could be easily neglected because was absent in existing models that time.

With this background, the communication and mediation of information between public agents and organizations can be studied through an analysis of the official acts published in official communications like D.O.U. The model presented in figure 02 aims to show how the communication of information between public agents happen in an informal dimension, how this communication evolves until it breaks through the edge between the informal and formal and how this interaction finally is manifested in an publication at official media.

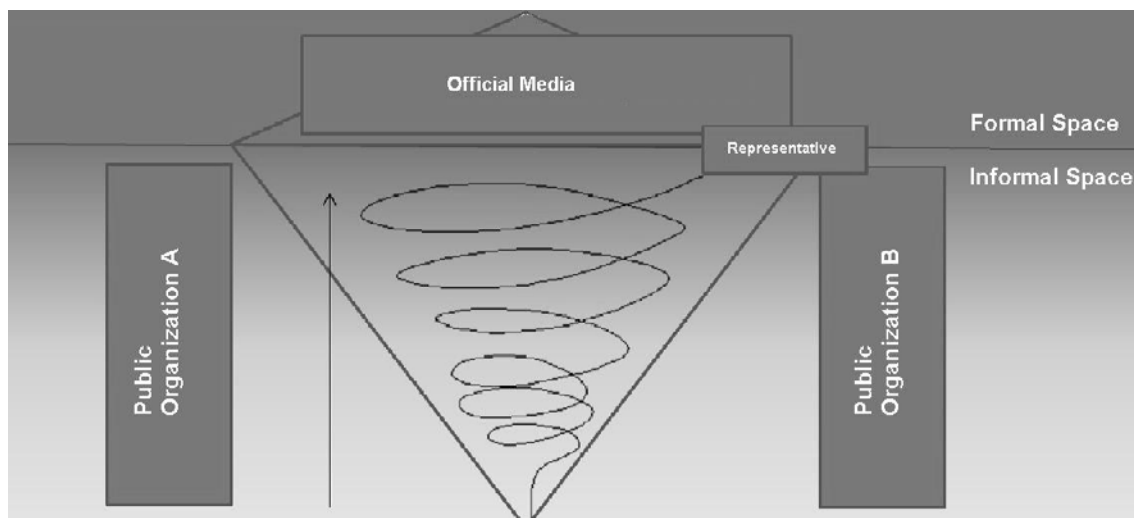


Figure 02 – public agents and organization information communication model

There is a mediation role played by the organization that prepares and publishes the decision. The elaboration process of the text to be sent to official media makes this player the representative of the group, who has the power of decide what is communicated and how.

The diamond shape in the back of the image represents this mediation phenomenon. The analogy to an iceberg intends to illustrate the fact that there are more information generated and a more intense communication flow going on before the decision in published in official media. And this publication only shows one piece of what really happened before. This understanding allows us to expand the analysis and infer that the evidence identified in official media represents a much more intense and significant interaction that precedes this evidence.

A pseudo-algorithm for building such network would be:

- For each page of the gazette:
 - Identify publication boundaries;
 - Identify person names inside the text.
 - For each pair of person found, establish a relationship.

Data was collected from the D.O.U (“Diário Oficial da União” – Federal Official Gazette), available online at <http://portal.imprensa nacional.gov.br/>, from 2012-january-01 to 2013-April-01, using an automated software solution developed for this task. This software downloads data, finds names using heuristics and produces files representing the networks. The software was developed using the methodology described by Soares (2014) and authors can promptly provide collected data upon request.

The following table shows data about the resulting network.

Table 02 – Numbers about the network built according to the methodology

Period analyzed	2012-january-01 to 2013-April-01
Documents analyzed	About 2100 files. (2.5 GB of text data)
D.O.U Sections	All 3: Acts, Human Resources and Contracts
Full Network size	Vertices: 221392, Edges: 11699909

3.2 Ego Networks

Ego networks are especially useful to address the position of an individual among other actors and his opportunities to exercise power raised from his structural location in the network and to measure his social capital provided by this position and the social support one can receive and offer in each network setup.

The hypothesis investigated in this paper is strongly related to power and hierarchy. Ego networks are a very useful tool to study hierarchical phenomena and the opportunities actors have to “escape” from formal restrictions in influence and information mediation.

In that sense, Cross and Cummings (2003) offer a very interesting point of view of the use of ego networks in studies in individual performance:

Research assessing individual (or egocentric) networks has established links between performance and structural holes, defined in terms of lack of connectivity among people in one’s network. People with networks rich in structural holes are more likely to be promoted early, enjoy greater career mobility and adapt to changing environments more successfully.

4. Analysis of the information flow and mediation in President Dilma Rousseff’s ego network

Mediation in Social Network Analysis can be seen as a phenomenon that makes “a person with a lot of friends and contacts have much better chances to get help or information. So, social ties are a measure of social capital, an asset to be used for

advantage” (NOOY, MVAR e BATAGELJ, 2005, p.138). Social place in the structure of a network can allow one to mediate communication between others, which is an opportunity to play relevant roles in the network leading, eventually, to benefits just because of the structural position.

In this section, we show how information mediation in a network of public agents including President Rousseff can be studied to identify prominent agents and patterns on communication.

4.1. Network definition and manipulation

We studied the information flow in ego networks using a graph mapped through the identification of all person names in the D.O.U. Each full name, identified using the methodology described above, was considered a node in the network. Relationships between these nodes were discarded if their number of occurrence were less than eight, which is the mean of the values of lines with more than 1 occurrence. This decision was enough to discard sporadic connections. Examples of such connections are listings of granted retirements or projects receiving government funds. In situations like these the co-citation of names is circumstantial, do not fit the “public agents and organization information communication model” presented in this paper and is unlikely to be found again, being, therefore, disposable.

Table 02 below shows the distribution of the connections by their number of occurrence in the network of agents whose name was found by the text processing method applied. It is clear that the line disposal policy suggested is not only feasible, but necessary to make the dataset useful, since the initial amount of connections identified, mostly with only one occurrence, is too high.

The agent chosen as central for the analysis of the network of agents participating in public Administration of Brazil was President Dilma Rousseff. The hypothesis is that around such an important actor would be raised a network where the communication flow would be intense enough to create evidence in the D.O.U. in the way described by the “public agents and organization information communication model” presented before.

Table 02 - Distribution of the connections by their number of occurrence

Line Values		Frequency	Freq%	CumFreq	CumFreq%
Lowest value of line: 1.00000000					
Highest value of line: 124.00000000					
(... 1.0000]	9831124	93.7047	9831124	93.7047
(1.0000 ... 42.0000]	660173	6.2924	10491297	99.9971
(42.0000 ... 83.0000]	279	0.0027	10491576	99.9998
(83.0000 ... 124.0000]	23	0.0002	10491599	100.0000
Total		10491599	100.0000		

So, the network, after removal of the low value lines, was reduced to a network containing only the nodes connected to Rousseff (the ego) directly or through one other actor. In other words, we created a network with the *2-neighbors* of the ego and the connections between this subset of actors: The ego (gray node in the middle), her first

level neighbors (white nodes) and her second level neighbors (black), as seen in figure 04 below.

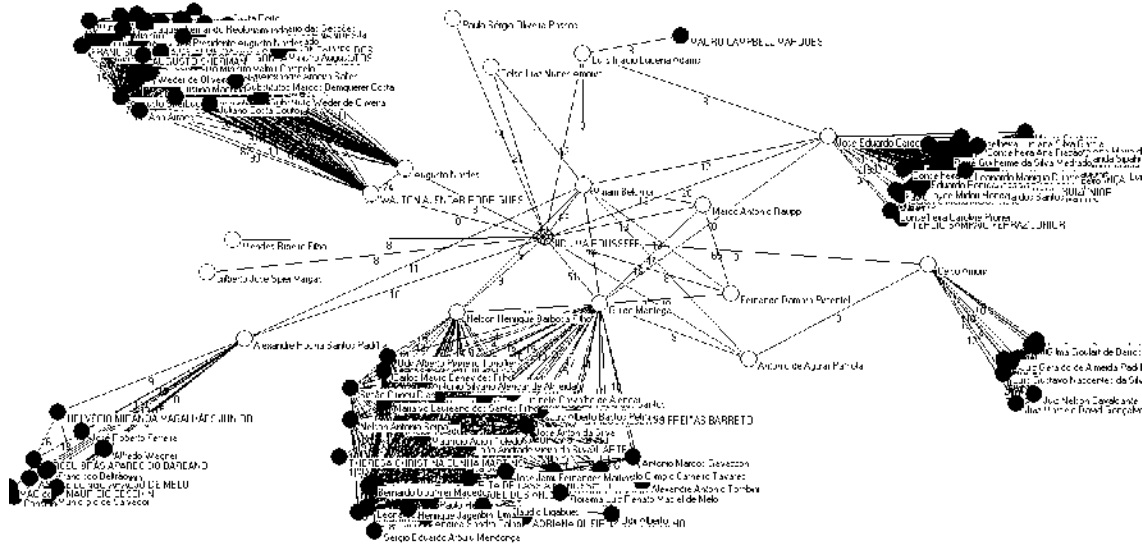


Figure 04 - President Dilma Rousseff's ego 2-neighbor network

4.2. Structural metrics and visual inspection

Visual inspection allows us to see a clear segregation between agents on the second level (blacks), who, structurally, are connected, at most, to two level one (white) individuals. There are five groups of agents on the second level, three of them connect to only one agent of level one and two of them connect to two agents of level one. There is a sixth “group” but it has only one member.

We can assume that these actors between the ego and the second level act as mediators of the interaction of those with the President. It comes from the way this view was built. A second level actor that eventually could have a closer access to the president would be shown as a first level one. We suppose this is what happened to “Nelson Henrique Barbosa Filho”, who is hierarchically below “Guido Mantega” but structurally at the same level.

Two agents in two groups played the mediator role. Nelson and Guido were Secretary and Minister of Economy respectively. Augusto Nardes and Walton Rodrigues are ministers of the Federal Court of Accounts (TCU – Tribunal de Contas da União). We can see in the network that they are ahead of the biggest groups of agents, where the information flow is probably bigger and demands more than one mediator. Yet, in the case of TCU it is interesting to see that Nardes and Rodrigues are still prominent for there are other seven ministers who are at the same top hierarchical level and do not seem to be as connected to the President as they are.

In this scenario, one interesting hypothesis to be cleared is that communication with the President is extremely restrict, since in the two level network only 8,45% (16 agents) have direct access to Rousseff. The fact that a second level is made by only 172 agents, a very small fraction of the 221387 agents in the original network reinforces this theory.

Finally, a structural aspect of this network should be highlighted: The clique formed by Mirian Belchior, Marco Antônio Raupp, Fernando Damata Pimentel, Guido Mantega e Dilma Rousseff. Cliques in social networks mean strong cohesion between the members what can be an indicator of a high level of cooperation, coordination and information exchange between these agents and the organizations they are ahead of. Identifying what causes this higher interaction requires further investigation. It can be either related to the mission of the organization, the personal affinity of these agents, party related issues or some other reason.

4.3. Information flow in Rousseff's neighborhood

Figure 05 shows a variation of the network shown in Figure 04 discussed in last section. To study dyadic restrictions, aggregated restrictions and structural holes in the ego network of President Rousseff, we removed all agents not connected directly to Dilma.

Generally speaking, dyadic restriction of a connection over an actor measures the opportunity this actor has to play a mediator role for others and the risk of having his communication to others mediated in case he loses this connection. As a rule of thumb, the bigger the restriction is the less mediation opportunities an actor has and more worried he should be about losing that connection.

This interpretation of the dyadic restriction metric, however, do not feels right to an ego network of a President. After all, since we used an open data resource to map these relationships, we assume the information flow between these agents is not secret, but the opposite: These are important information for cooperation and coordination of individuals involved, as discussed when the “public agents and organization information communication model” was shown in this paper.

What is expected, then, is more communication between agents in the Administration so the work would be done more efficiently, and from the President's point of view, there should be no interest in mediating information flow. In fact, big centralization of communication in the President's cabinet would represent an unnecessary overload, elevating risks of noise, delaying decisions and inflating bureaucracy.

So, the interpretation of this metric should be the inverse of what the literature often do, that is, from the President's point of view, a low restriction points out to a situation where the central power would be playing a mediator role which should be avoided. On the other hand, a high dyadic restriction mean low dependency of the other agents to the President, indicating a straight flow of information between them, situation that would lead to a better communication and should be pursued.

Table 03 shows values for dyadic restriction relative to the President and the lines between each agent and other neighbors. The positive correlation between these two metrics is expected and reinforces the notion that the more communication an agent has with others without the mediation of the President, bigger is the relative President's dyadic restriction, indicating more freedom for the agent in the network and less intervention of the central power in the information flow. We can see, for example, that Miriam Belchior has more connections with peers than any other, therefore, shows the biggest restriction to the ego.

Table 03 – Mediation metrics in the ego network

Agent	Dyadic restriction to President	Lines between other neighbors
Paulo Sergio Oliveira Passos	0,002222	0
Augusto Nardes	0,002626	2
Celso Luiz Nunes Amorim	0,006887	2
Marco Antonio Raupp	0,0147581	12
Miriam Belchior	0,1191908	32
Fernando Damata Pimentel	0,0177688	12
Guido Mantega	0,0873331	24
Alexandre Rocha Santos Padilha	0,0021115	2
Antonio de Aguiar Patriota	0,0138375	8
Celso Amorim	0,0017714	2
Jose Eduardo Cardoso	0,0096853	10
Gilberto Jose Spier Vargas	0,0007255	0
Luis Inacio Lucena Adams	0,0036674	6
Mendes Ribeiro Filho	0,0007255	0
Nelson Henrique Barbosa Filho	0,0029322	6
Walton Alencar Rodrigues	0,002626	2

These numbers can be aggregated in two indicators, to make the analysis process simpler. The aggregated restriction calculated for the President, calculated summing up all restrictions associated with her, was 0,2889 and the egocentric density, calculated as the density of the network when the ego is excluded, was 0.1484.

Aggregated restriction and egocentric density are hard to interpret alone. However, they are very useful to compare networks with similar size. So, in order to assess the impact of a more intense communication between the neighbors of the President in the network, these same metrics were calculated for a sub-network where the identified clique (Mantega, Belchior, Raupp, Pimentel and Rousseff) was removed, leading to a less interconnected network and, therefore, more dependent on the President. This new network showed 0.1501 for aggregated restriction and 0.0416 for egocentric density.

The removal of only four agents of the network dropped the aggregated restriction in 50% and the egocentric density in almost 70%. Such sensitivity in these metrics show that the values are small, because a very small change in the number of actors made a big change in the metrics. If these values were big, small perturbations in the network wouldn't lead to big impact on these metrics.

These calculations and comparisons of aggregated restriction and egocentric density allow us to believe that the communication of information between agents of public Administration in Brazil are strongly centered on President Rousseff, what may point out to difficulties implement direct articulation between these agents. A hypothesis raises then: An increment on the communication between the agents who deal directly with the President would lead to a situation where the information would flow more smoothly, fast, with less noise, and the President would have more resources to deal with other issues.

4.4. Structural proximity to President and nomination for high budget positions.

We analyzed another sub-network built from the original network after discarding every actor who did not connected directly to the President at least three times according to the methodology described. It is interesting to see that all the agents in Rousseff's 1-

neighbor ego network are public agents in first or second class positions, namely ministers and secretaries

There are several positions with different levels and responsibilities that the President can, freely, nominate or drop as she wishes. Over this 1-neighbor ego network we studied the structural proximity to the President and its relation to the budget of the ministry assigned to each element. This is a very unique characteristic of President's ego network and allow us to investigate if the closer they are to Rousseff, the more powerful they could be.

We assume that the bigger is the budget greater is the power and influence the one who controls it have. Thus, being nominated to head a ministry with big budget is something desirable from the individual point of view.

The hypothesis is that being structurally close to the President would lead to better opportunities to reach and keep these positions.

Table 04 shows the budget for each actor in the 1-neighbor network. Some of them have R\$0,00 budget because they aren't the top authority of the ministry and, thus, do not have full control of the budget. These information were manually extracted from government sources, specially the website *planalto.gov.br*.

Figure 05 shows this network drawn using the budget of the ministry as the size of the vertices. Using kamada-kawai energy algorithm, fixing Rousseff in the middle and using line values as similarities, we could produce a visualization that would place the most connected elements to the President closer the her correspondent node. In other words, the closer an element is of Rousseff, the greater is the number of connections it has with her, since more weigh means more communication.

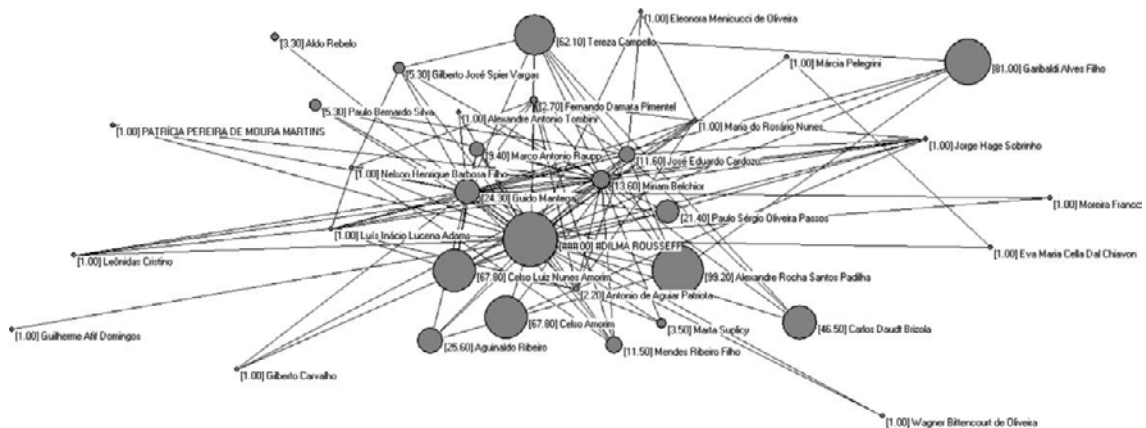


Figure 05 – President's 1-Neighbor network with proximity and time in position information

A visual inspection in this network leads us to believe that there really is a positive correlation between budget and number of connections with Dilma, because it seems that the bigger nodes are clustered together in the middle of the network, close to Rousseff, while the small ones are in the periphery of the net. Of course there are some exceptions in both groups.

Garibaldi Alves Filho, for instance, is represented by a big node far from Dilma. He was in charge of a ministry responsible for payments of retired workers in Brazil. So,

despite that his ministry has a huge budget, most of it is already designated. There is much less space for decisions that it seems.

To understand better the relation between the structural proximity to the President and the budget, we calculated the Pearson's correlation between these two numbers. It turned out that it was 0,6012, a statistically significant correlation that let us believe that, in fact, the elements structurally closer to the President are more likely to occupy more interesting positions in the Administration.

It is important to remember that such choices are political decisions, subject to lots of factors and influences and further additional research over this subject is essential to come to a stronger conclusion about the dynamics of power in politics.

5. Final comments

This paper makes an exploratory analysis of communications networks established around President Dilma Rousseff, based on data available at the federal official gazette. The structural metrics used were the agent's degree of relationships, dyadic restriction, aggregate restriction and egocentric density. Positive correlation coefficients indicate that the stronger the level of communication with the president, stronger is the tendency to being nominated for a position with large budget.

References

- COSTA, S. M. S. Comunicação para negócios: notas de aula da disciplina Fundamentos em Comunicação e Mediação da Informação. Programa de Pós-Graduação da Universidade de Brasília. 2013.
- COURTIAL, J.; LAW, J. A co-word study of artificial intelligence. *Social Studies of Science*, v. 19, n. 2, p. 301-311, maio 1989. Disponível em: <<http://www.jstor.org/stable/pdfplus/285145.pdf?acceptTC=true>>. Acesso em: 15 mar. 2016
- CROSS, R. CUMMINGS, J. N. Tie and Network Correlates of Individual Performance in Knowledge-Intensive Work. *Academy of Management Journal*. V. 47, n. 6, p. 928-937. December 2004
- DANCE, F. *Human communication theory: original essays*. Michigan: Holt, Rinehart and Winston, 1967.
- LE COADIC, Y. *A ciência da informação*. 2. ed. Brasília: Brinquet de Lemos, 2004.
- MCQUAIL, D.; WINDHAL, S. *Communication models for the study of mass communications*. Londres: Longman. 1993
- NOOY, W.; MRVAR, A.; BATAGELJ, V. *Exploratory network analysis with Pajek*. Cambridge: Cambridge University Press, 2005.
- SOARES, Rafael Henrique Santos. *Métodos para análise da comunicação e mediação da informação em organizações públicas por meio de redes sociais mapeadas a partir de publicações oficiais*. 2014. 155 f., il. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2014.