Social Network Analysis for Detecting Indicators of Favouritism in the Appointment of Corporate Managers

Luís A. M. Sales[1], André Ormastroni Victor[1], Rafael Basso[1], João F. C. Rocha[1], Bruno S. N. Contursi[1], Thiago Pereira Meirelles[1]

[1]Investigative Intelligence (IEMSI/II) – Petróleo Brasileiro S.A., Rio de Janeiro - RJ - Brasil.

{luis.sales, aovictor, rbasso, joao.felipe, bruno.contursi, thiago.meirelles}@petrobras.com.br

Abstract. *Several factors influence the choice of managers within organizations: previous experience, individual competencies, leadership attitudes, among many others. However, a central issue in the culture of corporate meritocracy is ensuring that there is no personal favouritism in the appointment of leading positions. This paper proposes a social network analytical approach to systematically measure the level of influence of connections between individuals in the appointments of managers. Promising preliminary results were validated through a case study containing historical data on workforce movement from a Brazilian mixed economy company.*

1. Introduction

Social Network Analysis (SNA) has been an analytical data approach applied in a variety of domains over the years [Tabassum *et al.* 2018]. Social network should not be understood strictly in terms of commercial social networks (Facebook, X, etc.), but encompasses any network of relationships. Some works have focused on applying SNA for fraud detection, where the main idea is to discover hidden relationships or other suspicious data patterns in the network [Victor *et al*. 2024, Akoglu *et al.* 2015, Amaral 2020, Lui *et al.* 2016, Paulo and Rodrigues 2020, Pourhabibi *et al.* 2020]. This paper proposes indicators to detect favouritism in the appointment of managers and executives within organizations, since meritocracy is a very important principle in modern corporate culture [Barbosa 1999], and the application of the proposed indicators to the concrete case of a Brazilian mixed economy company.

In this context, it is essential to discover if there are groups of vertices (employees) in the network formed by people and departments where the subordination relationships are preserved in the graph over time. In order to achieve this purpose, a set of metrics associated with the vertices and edges has been developed for revealing outlier cases, since their values deviate greatly from the mean value of other individuals in the network. This may be evidence that these groups have greater power of influence in the appointment of employees for leading positions [Safina 2015], even when organisational changing defines a new business unit hierarchy. Furthermore, this work can support initiatives for continuous preventive monitoring, so that if a suspicion of fraud or non-compliant action (e.g.: procurements, financial, etc.) occurs about an individual, the circle of managers with high functional connections can also be verified.

Section 2 of the paper presents the methodology used for the development of the work. Section 3 exposes analysis and interpretation of preliminary results obtained, while Section 4 forwards the final considerations and the future directions.

2. Methodology

This work is classified as applied research through a case study, since it involves applying SNA techniques on a real historical database of business function appointments of employees from a mixed economy company. Figure 1 illustrates the steps taken for the development of the case study.

Figure 1. Study case development method adopted

The first step concerns the graph modelling and definition of metrics that will identify the vertices with high level of influence in the network. The property graph scheme was built considering the network of relationships formed by the historical managerial appointments of the workforce, as illustrated by Figure 2. When an employee (Employee node, purple colour in the graph below) is appointed to a leading position (Manager, Coordinator, Supervisor, etc.) in a department (Department node, in yellow), a relationship is created between these nodes (APPOINTED_IN relationship, in red). Every relationship has a pair of properties (start date and end date) that describes the period in which the employee performed that position at the department. The departments form hierarchies (COMPOSED_OF relationship, in blue) which create an implicit subordination relationship between the higher-level managers with the lower-level managers at the hierarchy.

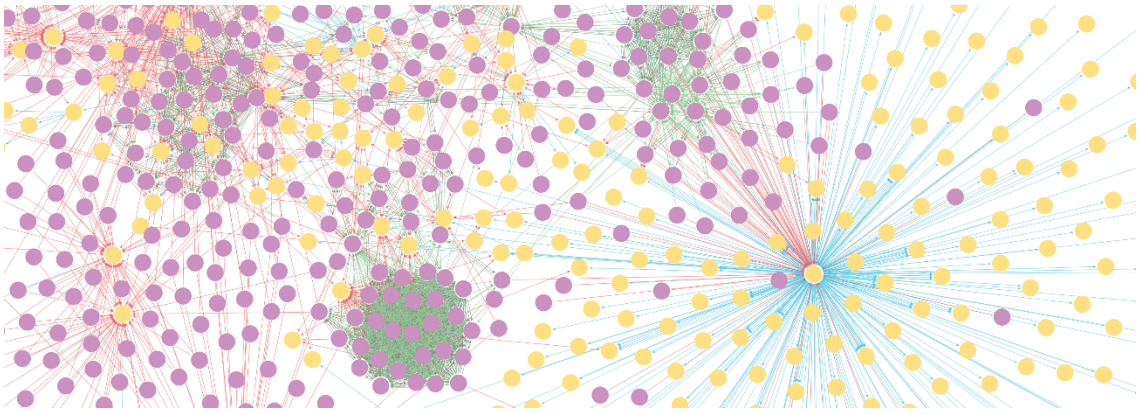

Figure 2. Graph subsample

With respect to metrics, the following indicators were developed to identify the employee nodes with higher level of influence in the appointment network:

- *$ID_{i,j}(D)\ \Delta t$ – Influence degree (IG) of node i to node j in the time period Δt in the context of department D (INFLUENCE_DEGREE edge, black colour in the graph above):* Consider a node *i* that started performing a position at the

department D on d_{start}. A relationship INFLUENCE_DEGREE is created between i and j if the node j was also allocated to D within Δt days after d_{start}.

- *$TID_i(D)$ Δt – Total influence degree (TIG) of node i in the time period Δt:* While the previous indicator captures the influence degree (*IG*) between *i* and *j* in the context of the department *D*, the total influence degree (*TID*) is the sum of the influence degrees in each department where i and j worked together and the metric $ID_{i,j}(D)$ Δt was calculated, for all j.
- *$L_{i,j}$ – Link of node i to node j*: This relationship (LINK, in green) captures the number of occurrences the nodes *i* and *j* worked together at the same department.

For each of the defined metrics, cases of favouritism are expected to be found in outliers, since it would be highly unlikely that, by pure chance [Aggarwal 2017], two unrelated employees were to start working with few days of difference between them (influence degree) or on the same department (link) multiple times. It should be noted, however, that false positives may also be captured by this methodology. For instance, other than by random, hierarchically superior managers that call the same consultants to work with them whenever they move in the company structure, not because of favouritism, but due to their proven competency. In order to distinguish such legitimate occurrences from mere favouritism, thorough inquiries should be conducted, resorting to methods such as interviews with the suspected persons, which falls outside the scope of the present paper.

Data extraction, transformation and loading were performed on corporative databases to shape originally tabular data into the proposed property graph. The next section presents the analysis of preliminary results obtained with the defined metrics.

3. Preliminary results

Due to the sensitivity of the data presented, the presentation of preliminary results is constrained by the Brazilian laws of privacy. As such, it will not be possible to show raw data and aggregated graphs will be the preferred means of visualisation, with min-max normalisation. However, the findings are still valid, since outliers exhibit an unusual behaviour in comparison to the inliners and are easily recognisable in such graphs, as extreme values usually concentrate on the tails of distributions.

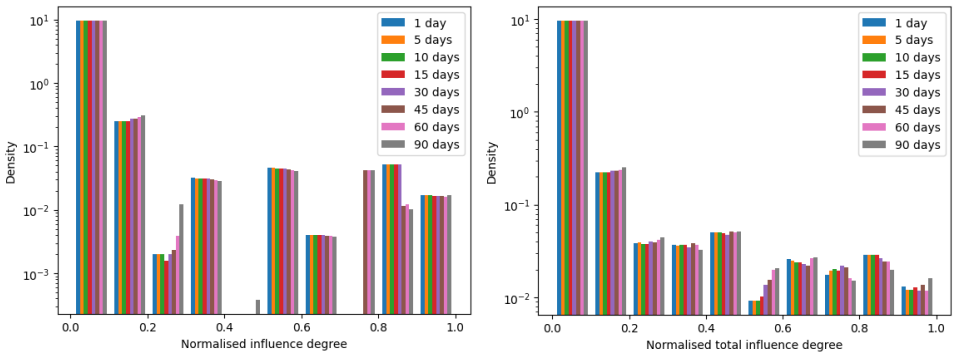

Figure 3. Normalised influence degree histogram (left) and normalised total influence degree histogram (right)

It is observed a general downward trend in the distribution of the proposed metrics (Figures 3 and 4, left), with many cases with low influence or link but few cases with high metric values. It is a desirable behaviour, since it allows identification of interesting cases. However, it is more remarkable for total influence and link metrics, which suggests that the influence degree exhibits less extreme behaviour. That is, it is easier to identify node (total influence degree) than edge (influence degree) outliers. Moreover, it would be expected that as the time interval increased, the influence degree would accompany it. However, it was not observed homogeneously, implying that for most cases employees start working in the same department with few days of difference.

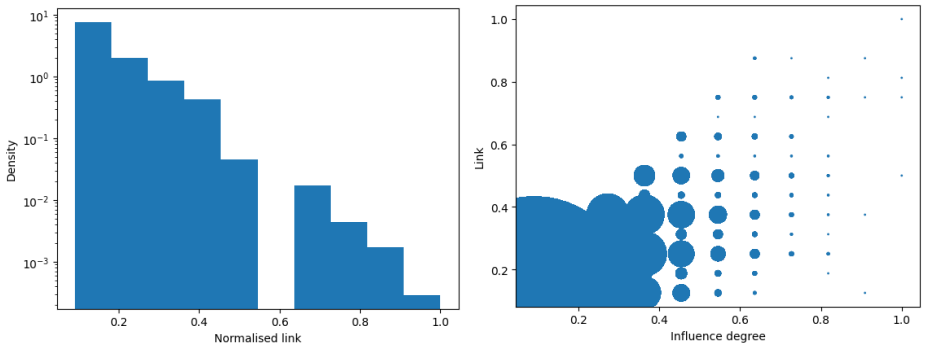

Figure 4. Link histogram (left) and Influence degree (30 days) and link scatter plot (right)

The scatter plot on Figure 4 (right) shows dot size proportional to the frequency of the link and influence degree (30 days time window) data. This graph shows a positive correlation between the influence degree and the link between two managers, however it is not 1. There is a correspondence between the phenomena that both indicators are measuring, but they are still different metrics. On its own, each metric generates its extreme values, however even more clear outliers can be found by combining both metrics, the right top points, which should be prioritised in further investigations.

4. Final considerations

This work presented a study case where the historical base of appointments of leading positions was represented as a directed graph. Through SNA techniques, we analyse nodes and edges from network to identify groups of employees that preserve their subordination relationships over time. Metrics associated with the graph were calculated and demonstrates a certain level of influence some employees have in the appointment of other managers, which may be evidence of favouritism.

Despite the initial stage of development, the preliminary results identified a small set of outlier cases where, in fact, employees have recurrent subordination relationships. Depending on the hierarchical structure surrounding the position to be appointed, we would be able to identify a small group of people with a higher chance of being appointed than others. As future developments for this work, we intend to develop new metrics and segment the descriptive analyses according to each business unit. In addition, we intend to combine the developed metrics (influence degree and link) with other classic SNA metrics (e.g. degrees, betweenness, closeness, etc) and develop graphical views of the main results and cases of interest to expose them to a more executive audience within the organization.

References

Aggarwal, C. C. (2017). *Outlier analysis*. Springer International Publishing.

Akoglu, L., Tong, H., & Koutra, D. (2015). Graph based anomaly detection and description: a survey. *Data mining and knowledge discovery*, *29*, 626-688.

Amaral, W.S. (2020). Análise de Grafos para Apoio em Auditoria de Licitações Públicas. https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=103180471-12.

Barbosa, L., 1999. Igualdade e meritocracia: a ética do desempenho nas sociedades modernas. Fundação Getúlio Vargas Editora, Rio de Janeiro.

Liu, J., Bier, E., Wilson, A., Guerra-Gomez, J. A., Honda, T., Sricharan, K., Gilpin, L., & Davies, D. (2016) Graph Analysis for Detecting Fraud, Waste, and Abuse in Healthcare Data. AI Magazine, 37(2), p. 33-46.

Paulo, A.C.R.M. and Rodrigues, A.P.S.P. (2022) Visualização de relacionamentos utilizando grafos como ferramenta de fiscalização de recursos públicos. Academic Journal on Computing, Engineering and Applied Mathematics, 3(2), p. 1-12.

Pourhabibi, T., Ong, K. L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems*, *133*, 113303.

Safina, D. (2015). Favouritism and nepotism in an organization: Causes and effects. *Procedia economics and finance*, *23*, 630-634.

Tabassum, S., Pereira, F.S.F, Fernandes, S., Gama, J. (2018). Social network analysis: An overview. WIREs Data Mining and Knowledge Discovery, 8(5): e1256

Victor, A. O., Sales, L. A. M., Moreira, R. S., de Paula, T. S., de Moraes, C. E. C., Meirelles, T. P., Lima, L. G. G. B., Rocha, J. F. C. and Contursi, B. S. N. (2024) Detecção de Indícios de Fraude em Licitações através de Técnicas de Análise de Redes Sociais. In: Rio Oil & Gas Expo and Conference (in press).