



PRIMARY RESEARCH

# A methodology for strategic diagnosis of business corruption behaviours using network analysis

Dana Nedeia<sup>1\*</sup>, Ioan Petrișor<sup>2</sup><sup>1,2</sup> Faculty of Economics and Business Administration, West University of Timișoara, Timișoara, Romania

## Keywords

Business corruption  
Behaviour  
Network analysis  
GDELT

**Received:** 7 October 2019**Accepted:** 8 November 2019**Published:** 23 December 2019

## Abstract

Business corruption behaviours directly impact the business environment and the way in which strategic diagnosis is conducted at organizational level. Thus, the diagnosis of these business corruption cases becomes highly important in the process of strategic decisions. The methodology we use is based on networks analysis. We propose a Strategic Diagnosis Based on Networks Analysis (SDBN) which incorporates modularity and clustering coefficient measurements. These measurements allow for a cross-diagnosis of business corruption behaviours. A large quantity of event data from mass-media was extracted from the Global Database of Events, Language, and Tone regarding European Union business corruption, over the year of 2019. The data encompasses countries and organisations & institutions as the nodes, and the number of times that <<European Union>> and <<business corruption>> appear together in our search, as the edges of the networks. These networks are then located in our proposed SDBN model, and are separated in concentration or diversification strategies, weakly connected or highly connected. The distinctiveness of our research consists of involving the theory of network analysis in the process of strategic diagnosis. The results reveal the specific business corruption behaviours conducted by actors involved in business corruption cases, with the limits being given by the usage of big-data and the magnitude of corruption behaviours. Future research will intend to apply the proposed SDBN to a larger sample of data and to different perspectives.

© 2019 The Author(s). Published by TAF Publishing.

## INTRODUCTION

This paper uses network analysis in order to propose a methodology for business corruption strategic diagnosis at European Union level, using data extracted from Global Knowledge Graph (GKG) Network Visualizer included in the Global Database of Events, Language, and Tone (GDELT) Project. We propose the construction of a strategic diagnosis model based on network analysis clustering and modularity measurements, named SDBN model. In this paper we intend to give a broad use of this model in the process of strategic diagnosis of business corruption.

One of the major difficulties in studying corruption is given by the way in which corruption is defined, as this directly determines the way in which it is measured and modelled [Jain \(2001\)](#). In this paper we rely on defining the corruption as being the use of public power for a personal gain, considering the occurrence of corruption at the overlapping of

public power and private wealth ([Rose-Ackerman, 2013](#)). The above definition implies both supply and demand dimensions. [Dixit \(2016\)](#) considers that the focus in corruption studies is generally in the demand side, as most of the studies tend to focus on the government/public corruption, while the supply side is also worth the study, since private individuals or business organisations are the other major part of corruption transactions. Given this gap considered, our study takes into consideration the business corruption side, as the supply side of the corruption transactions. Many of the corruption studies are also focused on population perception of corruption by the population, while not many of them focuses on the mass media and press perception over corruption phenomenon. In this study we use this perception as a tool for the strategic diagnosis of business corruption, as mass media and press are known for constantly exposing the bribery of government and pri-

\*corresponding author: Dana Nedeia

†email: [dana.nedeia94@e-uvt.ro](mailto:dana.nedeia94@e-uvt.ro)

vate authorities as well as business corruption scandals (Rose-Ackerman, 2007).

The practical implications of our paper are to be considered by both public and business sectors, as they can prove to be useful in inspiring the strategic directions needed in corrupt environments.

The proposed methodology can be used by business organizations in order to conduct the diagnosis of business corruption strategies followed in different geographic areas and specific business sectors. Our proposed methodology can be also used by international and national institutions in order to construct public policies for the anti-corruption fight.

The paper structure follows a literature review introducing the main theoretical topics discussed, the methodology and data, which introduces the strategic diagnosis model, and the way of collecting and analysing the data, followed by the results and conclusion of the study among with specific discussions and future research directions.

## LITERATURE REVIEW

The term corruption is often used to refer to a wide range of illegal or illicit activities (Morgan, 1998). Among literature review there is a consensus on defining corruption as the use of public power for a personal gain, considering the occurrence of corruption at the overlapping of public power and private wealth (Haass & Ottmann, 2017; Rose-Ackerman, 2007).

From this definition, two approaches derive. The one in which the corruption transaction is initiated by the public official, and the other, which is initiated by the private individual or business organizations.

A common corruption transaction is one in which a private individual or a business organization pays a public official in return for a benefit (Rose-Ackerman, 2007). This transaction has both supply and demand dimensions. Dixit (2016) states that the emphasis in regard to corruption studies is mainly in the demand side, as studies tend to focus on government corruption, while the supply side should also be studied, since the business organizations are a major part of corruption transactions.

From this view, we consider the business corruption phenomenon as being the behaviour of business organizations in the corruption transactions, linking it with other studies that refer to this behaviour as being organizational corruption (Hogdson & Jiang, 2007).

Previous studies relating mass media and corruption intend to detect specific relations between mass media and corruption actions. The investigation over implications of

mass media and mass media freedom to corruption (Camaj, 2013; Suphachalasai, 2005), the way in which mass media promotes accountability in nondemocratic countries (Enikolopov, Petrova, & Sonin, 2018), or the effects of mass media over public policies (Strömberg, 2001) are some of the studies using mass media as an analysis tool in regard to corruption.

We observe a need to centre the study of corruption to mass media and press in order to the mass media perception over corruption, and the status quo of the corruption phenomenon at a given time.

Johnston (1996) considers that corruption should be treated as a systemic problem, in both an analytical sense and in the sense of difficulty to address the issue, common to all societies.

In our study, we consider the relation between mass media and business corruption to be a strategic diagnosis tool. The hypothesis of our study is that, by using the perception of mass media over corruption, a diagnosis model can be constructed and applied, determining specific business corruption strategic areas.

For the purpose of this study, we conduct our research on business corruption, as being considered the supply side of a corruption transaction.

## METHODOLOGY AND DATA

This paper introduces a methodology for strategic diagnosis of business corruption behaviours using network analysis. We propose the construction of a model based on networks analysis measurements, named SDBN model. We developed this model previously using Global, European, EU and Romanian data for business corruption and also corruption (considered as public corruption) (Petrişor & Nedeia, 2019). Also, we have previously used this model in order to detect the business corruption strategies followed at Romanian level (Nedeia & Petrişor, 2019).

Network analysis uses graph theory in order to construct the networks which are formed of nodes connected by edges (Barabási & Bonabeau, 2003; Gephi, 2016). One important advantage of the usage of network analysis is given by the fact that the topological properties are universal, and they have cognitive potential to describe almost any entity (Barabási & Bonabeau, 2003).

Some of the most important measures used in this paper are modularity, average clustering coefficient, average degree and average path length. Modularity is used in order to extract the community structure of large networks, as it measures the strength of division of a specific network into communities or clusters (Blondel, Guillaume, Lambiotte, &

Lefebvre, 2008). The clustering coefficient measures the degree to which nodes in a graph tend to cluster together in a community. The average clustering coefficient is the mean of all clustering coefficients, being an indicator of how complete the neighbourhood of a node is (Watts & Strogatz, 1998).

The degree of a node calculates the number of edges connected to that specific node. The average degree calculates the mean of all the degrees, the result being useful in observing the nodes with higher connections in the network (Barabási & Bonabeau, 2003).

Average path length is defined as the average number of steps needed for all the nodes in the network to reach each other (Barabási & Bonabeau, 2003).

In order to collect the data, we use GKG Network Visualizer included in the GDELT Project. The Global Data on Events, Location, and Tone (Leetaru & Schrodt, 2013) project, supported by Google, covers hundreds of millions of event records, which are extracted from broadcast, print, and online news sources from all over the world in real time, being updated at every 15 minutes (Leetaru & Schrodt, 2013).

In our research the year of 2019 is the reference time interval. GKG Network Visualizer is scanned for the research parameters of, "European Union" and, "business corruption" for each month separately, and a list of all organizations and locations (countries) is compiled as the nodes of the network and the number of times that our search parameters occur together as the edges of the networks.

Downloaded data is the compiled and analysed using Gephi Software. For each parameter we calculate the modularity coefficient, average clustering coefficient, the average degree and the average path length.

A high modularity class with more communities indicates dense connections between the nodes within communities and sparse connections between nodes in different communities.

A low modularity class implies sparse connections between the nodes within communities and dense connections between nodes in different modules.

In order to construct the SDBN, we determined the communities in each network following the community detection model introduced by (Blondel et al., 2008) and based on modularity and average clustering coefficient we con-

structed a diagnosis model (Figure 1).

AVERAGE CLUSTERING COEFFICIENT	HIGH	Concentration strategies (highly connected)	Diversification strategies (highly connected)
	LOW	Concentration strategies (weakly connected)	Diversification strategies (weakly connected)
		LOW	HIGH
		MODULARITY CLASS	

**FIGURE 1.** Strategic diagnosis based on networks theoretical model

The theoretical SDBN generated is then applied on each scale and parameter in order to determine the possible corruption strategies followed at EU level by different stakeholders (countries and organizations & institutions). The model helps in the specific diagnosis of business corruption context. In order to determine the low and medium intervals we calculated the median of the results, from both modularity class and clustering coefficient. These measurements have values in the interval [0-1].

The main characteristic for the concentration corruption strategies is given by the low modularity class. This means that the strategies conducted in this area are focused on constructing few communities, with sparse connections inside each community and dense connections between the different communities in the network. In this communities the main focus is on the relations between different communities rather than on the inside communities' relations. For the diversification corruption strategies, the main characteristic is the low modularity class with dense connections inside communities. These dense connections allow the formation of small communities which are strongly connected and are focused on corruption strategies inside the community, and not as much outside communities.

The downloaded data for the specified time intervals and scales present a large number of nodes and edges which allows for a relevant network analysis, as shown in Table 1.

**TABLE 1.** Number of nodes and edges

2019	Nodes		Edges	
	Countries	Organizations & Institutions	Countries	Organizations & Institutions
January	171	241	5864	1376
February	178	307	6259	2036
March	186	522	7508	6026
April	174	285	6408	1666
May	154	314	4510	2513
June	168	256	5741	1534
July	169	295	6960	2063
August	153	159	3370	845
September	160	263	3916	1598
October	163	402	5432	3106
November	184	703	8158	4345
December	183	523	4942	3180

## RESULTS

The networks measurements are presented in Table 2. The results reveal a difference between countries networks and organizations & institutions networks, as most of the countries networks present a lower modularity class with less communities, while most of the organizations & institutions present a higher modularity class. These results are similar with our previous research in terms of strategic diagnosis of corruption and business corruption using network analysis (Nedeia & Petrișor, 2019; Petrișor & Nedeia, 2019).

The average degree results show a high number of edges

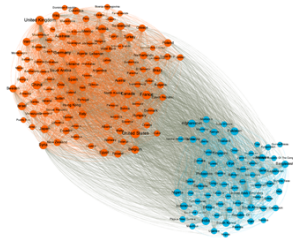
connected to the nodes in the countries network (from 44,052 to 88,674), indicating that, on average 44-88 edges are connected to a node in these networks, while for organizations & institutions networks, the average degree is lower (from 10,629 to 23,088), indicating on average, that 10-20 edges are connected to a node in these networks. As for the countries networks the average path length is lower than for the organizations & institutions. Still, for both levels, this is close to 2, meaning that these networks are composed from nodes that reach each other on average by passing to 2 other nodes.

**TABLE 2.** Network measurements

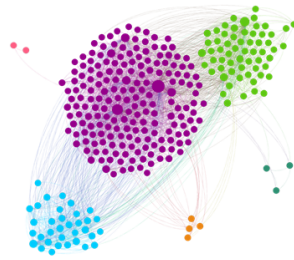
2019	Modularity/(Communities number)		Average Clustering Coefficient		Average Degree		Average Path Length	
	Countries	Organizations & Institutions	Countries	Organizations & Institutions	Countries	Organizations & Institutions	Countries	Organizations & Institutions
January	0,072/3	0,3/5	0,809	0,862	68,585	11,419	1,598	1,967
February	0,124/3	0,502/5	0,814	0,862	70,326	13,264	1,603	1,98
March	0,155/3	0,511/6	0,810	0,862	80,731	23,088	1,564	1,968
April	0,142/4	0,172/9	0,809	0,856	73,655	11,691	1,574	1,965
May	0,161/3	0,402/9	0,803	0,886	58,571	16,006	1,617	1,961
June	0,098/4	0,251/8	0,796	0,858	68,345	11,984	1,591	1,969
July	0,077/3	0,207/6	0,838	0,854	82,367	13,986	1,510	1,973
August	0,075/5	0,11/6	0,766	0,862	44,052	10,629	1,715	1,949
September	0,079/3	0,168/7	0,784	0,846	48,95	12,152	1,69	1,982
October	0,11/3	0,181/7	0,799	0,858	66,65	15,453	1,58	2,018
November	0,141/2	0,137/10	0,853	0,907	88,674	12,361	1,52	2,087
December	0,062/2	0,322/10	0,804	0,894	54,011	12,161	1,70	2,021



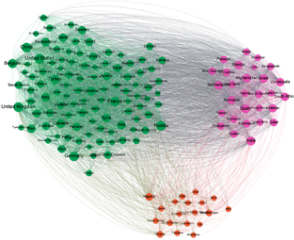
**FIGURE 2.** Organizations & institutions: August 2019 (Concentration strategies highly connected)



**FIGURE 3.** Organizations & institutions: July 2019 (Diversification strategies highly connected)

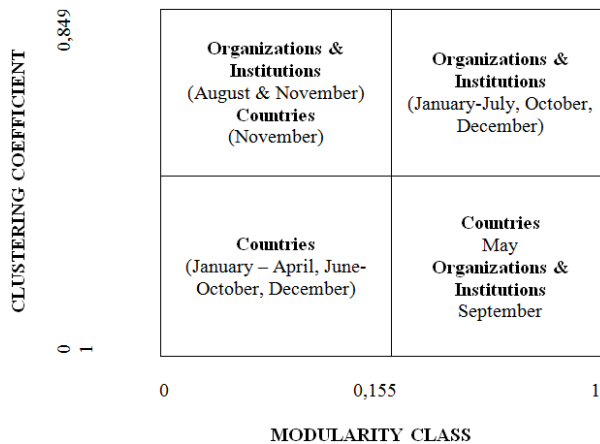


**FIGURE 4.** Countries network: December 2019 (Concentration strategies weakly connected)



**FIGURE 5.** Countries network: May 2019 (Diversification strategies weakly connected)

Figures 2, 3, 4, 5 are illustrated below in order to exemplify the networks and the communities resulted in terms of countries and organizations & institutions, one of each area.



**FIGURE 6.** Strategic diagnosis based on networks model

Figure 6 presents the results in terms of the strategic diagnosis of business corruption at EU level for the entire year of 2019. Based on our proposed methodology we identify the specific strategies followed by the stakeholders in the network (countries and organizations & institutions). We detect a majority of countries networks in the area of concentration strategies (weakly connected), with the exception of the month of November which is detected by our model in the area of concentration strategies (highly connected) and the month of May which is detected to be in the area of diversification strategies (weakly connected). As for the organizations & institutions networks, the majority of them are

in the area of diversification strategies (highly connected) with the exception of the month of September which is located in the area of diversification strategies (weakly connected) and the months of August and November which are located in the area of concentration strategies (highly connected).

**DISCUSSION AND CONCLUSION**

The methodology proposed in this article indicates the specific business corruption strategies followed by actors involved in business corruption discussions around mass-media in the year of 2019. We detect these strategies for each month, in order to construct and validate the proposed SDBN model. For this, the stakeholders taken into consideration are countries and organizations & institutions, as actors involved in business corruption discussions. Each specific area from the model indicates different strategies followed by these actors.

As for the countries networks, the specific strategies are mostly in the area of concentration strategies (weakly connected), which indicates a lower number of communities and lower clustering power, given by the low average clustering coefficient. This implies that these strategies are focused on constructing a few communities which have dense connections between them, but sparse connections inside each community. The dense connections between communities are a result of international relations between countries, while the sparse connections inside each community are a result of the magnitude of these actors, being difficult to maintain strong relations inside communities.

The month of May is detected by our model as being in the area of diversification strategies (weakly connected). This month coincides with EU Parliament Elections, and the diversification strategies may be explained by this, as previous studies shown that during some intense political events, such as elections, corruption discussions tend to be more pronounced (Rose-Ackerman, 2007).

The month of November is detected in the area of concentration strategies (highly connected) with a relatively high average clustering coefficient. In order to specify the reasons behind this raise in the power of countries to cluster together, future studies are needed, and we intend this to be one of the future research directions.

The countries networks are analysed in order to detect and diagnose the macro-environment, being useful in the analysis of the organizations & institutions networks, as these are our main focus on this study, consider the actors involved in the supply side of business corruption transactions.

The organizations & institutions networks are mostly posi-

tioned in the area of diversification strategies (highly connected), indicating a large number of communities that also have dense connections inside communities, but sparse connections between the communities in the networks. These allow the formation of small communities that are strongly connected and focused on corruption strategies inside the community, with the neighbourhood communities being not as related between them. These strategies are a result of these organizations & institutions being highly related to each other in the business corruption discussions around mass media.

The months of August and November are located in the area of concentration strategies (highly connected) and the month of September which is located in the area of diversification strategies (weakly connected), these results need some future research in terms of micro-environment diagnosis and specific business corruption cases discussed in these months, as it is clear from the present results that these month show some specific discussions that are not following the trend.

The analysis in this article brings a new validation to our SDBN model, as it helps detecting the specific business corruption behaviours of organizations & institutions and countries, followed during the year of 2019. This analysis indicates differences between countries networks and organizations & institutions networks in terms of communities number as well as clustering power.

These results conclude with the diagnosis of business corruption strategies followed by actors involved in the mass media discussions regarding the topic of business corruption at EU level, both for countries and for organizations &

institutions. The detected strategies are useful in the process of business strategy diagnosis, as well as in the process of constructing anti-corruption public policies, as the strategies detected are of actuality and present the situation in the recent year. Our proposed model may be applied in real time, as the data source we propose (GDELT) updates the data at every 15 minutes.

The limitations of our research are given by the still early usage and validation of our SDBN model, as well as the usage of big data and by the magnitude of business corruption behaviours, as well as the fact that mass media covers a corruption subject only from one perspective, and this perspective may differ from the public perspective. Another limitation of our research is the usage of organizations & institutions parameter, as GDELT currently offers no option in separating these. A way in which this limitation can be overcome is the usage of raw data, which is part of a future research.

Future work intends to apply the SDBN to a larger sample of data with the usage of dynamic networks and time intervals data. Future research work also needs to involve the application of this diagnosis model at the level of multiple countries and to also extend its application to some specific business sectors, and to validate it by comparing this perspective with other perspectives on business corruption, as the public/population perspective. Another future research work will include the specific diagnosis for those months that we detected as not following the trend, in which cases some specific business corruption cases may have been discussed in the mass media which are worth further attention.

## REFERENCES

- Barabási, A., & Bonabeau, A.-L. (2003). Scale-free networks. *Scientific American*, 288(5), 60-69. doi:<https://doi.org/10.1103/revmodphys.74.47>
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: Theory and experiment*, 20(10), 10-30. doi:<https://doi.org/10.1088/1742-5468/2008/10/p10008>
- Camaj, L. (2013). The media's role in fighting corruption: Media effects on governmental accountability. *The International Journal of Press/Politics*, 18(1), 21-42. doi:<https://doi.org/10.1177/1940161212462741>
- Dixit, A. K. (2016). Corruption: Supply-side and demand-side solutions. In *Development in India*. New Dehli, India: Springer.
- Enikolopov, R., Petrova, M., & Sonin, K. (2018). Social media and corruption. *American Economic Journal: Applied Economics*, 10(1), 150-74. doi:<https://doi.org/10.1257/app.20160089>
- Gephi. (2016). *The open graph viz platform*. Retrieved from <https://bit.ly/2xGmeSo>
- Haass, F., & Ottmann, M. (2017). Profits from peace: The political economy of power-sharing and corruption. *World Development*, 99, 60-74. doi:<https://doi.org/10.1016/j.worlddev.2017.07.006>
- Hogdson, G. M., & Jiang, S. (2007). The economics of corruption and the corruption of economics: An institutionalist perspective. *Journal of Economic Issues*, 41(4), 1043-1061. doi:<https://doi.org/10.1080/00213624.2007.11507086>

- Jain, A. K. (2001). Corruption: A review. *Journal of Economic Surveys*, 15(1), 71-121. doi:<https://doi.org/10.1111/1467-6419.00133>
- Johnston, M. (1996). The search for definitions: The vitality of politics and the issue of corruption. *International Social Science Journal*, 48(149), 321-335. doi:<https://doi.org/10.1111/1468-2451.00035>
- Leetaru, K., & Schrodt, P. A. (2013). GDELT: Global data on events, location, and tone, 1979-2012. In *ISA annual Convention, California, CA*.
- Morgan, A. (1998). Corruption: Causes, consequences, and policy implications. In *The Conference on Asia Foundation, San Francisco, CA*.
- Nedeia, D., & Petrişor, I. (2019). Detecting business corruption strategic behaviours using network analysis. *Emerging Markets Economics and Business*, 4(7), 166-180.
- Petrişor, I., & Nedeia, D. (2019). Projecting a strategic diagnosis system of corruption based on networks analysis. In *Innovation for Sustainable Management and Entrepreneurship* Timisoara, Romania.
- Rose-Ackerman, S. (2007). *International handbook on the economics of corruption*. New York, NY: Edward Elgar Publishing.
- Rose-Ackerman, S. (2013). *Corruption: A study in political economy*. London, UK: Academic Press.
- Strömberg, D. (2001). Mass media and public policy. *European Economic Review*, 45(4-6), 652-663. doi:[https://doi.org/10.1016/s0014-2921\(01\)00106-4](https://doi.org/10.1016/s0014-2921(01)00106-4)
- Suphachalasai, S. (2005). Bureaucratic corruption and mass media. *Environmental Economy and Policy Research*, 55(7), 45-50. doi:<https://doi.org/10.2139/ssrn.722403>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(66), 440-450. doi:<https://doi.org/10.1038/30918>