

Uma Abordagem Baseada em Redes Complexas para Análise de Depoimentos Legais

Thais G. Almeida¹, Fabíola G. Nakamura¹, Eduardo F. Nakamura¹

¹Universidade Federal do Amazonas (UFAM) – Manaus, AM – Brasil

{tga, fabiola, nakamura}@icomp.ufam.edu.br

Abstract. *In this paper, we present a methodology to identify influential entities and important relationships in legal testimonies. First, we model the testimonies as complex networks that represent the social relationships among the people that are referred to in the testimonies. Then, we use: (1) centrality metrics (e.g., eigenvector, betweenness) to identify important entities; (2) community detection algorithms to find correlated groups (social nuclei). In addition, we propose the use of the Shannon entropy to quantify the discrimination level of the centrality metrics. As a case study, we modeled five public plea agreements of the Lava Jato Operation. The results show that the entities that we identify as influential have become targets of investigations. We declare that this work is not intended to be used as legal evidence nor to determine the level of fairness of the testimonies and investigations. We simply present a social analysis that is based on the topological structures of the networks that represent testimonies, whether or not those testimonies represent the truth.*

Resumo. *Neste trabalho, apresentamos uma abordagem para identificar entidades e relacionamentos influentes em depoimentos legais. Como estudo de caso, modelamos cinco delações públicas da operação Lava Jato como redes (complexas), a fim de representar as relações sociais entre as pessoas citadas nos depoimentos. Para identificar entidades importantes, utilizamos medidas de centralidade (e.g., eigenvector, betweenness) selecionadas com base na entropia de Shannon. Os relacionamentos influentes foram detectados pelo algoritmo de Louvain. Nossos resultados mostram que os nós (pessoas) identificados como influentes nas redes se tornaram alvo de investigações. Os autores declaram que não são objetivos deste trabalho servir de evidência legal ou determinar o grau de veracidade dos testemunhos e justiça das investigações. O trabalho simplesmente apresenta uma análise social baseada nas estruturas topológicas das redes que representam depoimentos, quer esses depoimentos representem a verdade ou não.*

1. Introdução

Redes complexas consistem de grafos cuja estrutura topológica é irregular, não trivial e evolui dinamicamente ao longo do tempo [Boccaletti et al. 2006]. Tais redes são utilizadas para representar e investigar diferentes cenários, como a disseminação de doenças [Brandão and Moro 2016], o comportamento social online [Kossinets and Watts 2009] e a criminalística [Rostami and Mondani 2015].

A partir de características das conexões entre entidades de uma rede complexa, é possível identificar personagens chaves do contexto considerado. Para tanto, podem ser

utilizadas medidas de centralidade (e.g., *betweenness*) e transitividade (e.g., *local clustering coefficient*). A ideia por trás dessas medidas é intuitiva, entretanto suas interpretações devem ser realizadas com cuidado, uma vez que nem sempre a medida escolhida é discriminativa para compreender uma rede. Se por um lado o alto grau de um nó sugere sua importância em uma rede, isto pode não ser observado em outras redes [Campana 2016].

Quanto ao cenário criminalístico, a análise de redes é considerada uma ferramenta chave da inteligência criminal desde os anos 1970 [Masys 2013]. Contudo, tal análise é desafiadora, pois: (i) as bases de dados em geral são limitadas, em razão da sensibilidade dos dados [Rostami and Mondani 2015]; (ii) relatórios criminais frequentemente apresentam contradições, erros de digitação e *bias* dos envolvidos [Morselli 2013]; e (iii) nós ou arestas ausentes podem levar a interpretações erradas sobre a rede [Campana 2016].

A Operação Lava Jato, deflagrada em 2014 pela Polícia Federal do Brasil, tem como objetivo investigar a prática de crimes financeiros e desvio de recursos públicos. Para incentivar criminosos a colaborarem com as investigações, são firmados acordos de leniência (delações premiadas), nos quais em troca de informações criminosos podem ter suas penas reduzidas ou até mesmo extintas¹. Nesse contexto, a identificação de entidades influentes é importante, uma vez que a partir de informações presentes nas delações premiadas, as autoridades policiais podem direcionar suas investigações. Além disso, o forte interesse da sociedade em relação à Operação Lava Jato e a disponibilidade pública de delações motivam a avaliação destes documentos.

As principais contribuições desse trabalho são: (i) uma abordagem de redes complexas para identificar entidades influentes por meio de medidas de centralidade; (ii) a utilização da entropia de *Shannon* para selecionar quais medidas de centralidade são mais discriminativas para determinar entidades influentes; e (iii) detecção de comunidades para verificar se os agrupamentos refletem alguma organização social real. Embora o contexto utilizado sejam as delações da operação Lava Jato, ressaltamos que os resultados encontrados refletem os relatos dos delatores, sejam estes verídicos ou não, tendenciosos ou não. É importante ressaltar que não são objetivos deste trabalho: (i) servir de evidência legal; (ii) determinar o grau de veracidade dos testemunhos; (iii) determinar o grau de justiça das investigações. O trabalho simplesmente apresenta uma análise social baseada nas estruturas topológicas das redes das delações em questão.

O restante deste trabalho está organizado como segue: a seção 2 discute trabalhos relacionados e a seção 3 descreve a metodologia utilizada. Os resultados são apresentados na seção 4 e a seção 5 descreve as conclusões, os desafios e as direções futuras.

2. Trabalhos Relacionados

Técnicas de análise de redes são utilizadas com o objetivo de oferecer novas perspectivas de como visualizar e interpretar dados criminais complexos. A detecção de comunidades em grafos, por exemplo, é uma técnica utilizada para um maior compreensão sobre a organização do sistema representado, bem como sobre sua função. Calderoni and Piccardi [2014] aplicaram métodos de análise de comunidade em uma rede contendo integrantes de famílias de máfias italianas, com objetivo de explorar a relevância de subgrupos em redes criminais. Para tanto, os autores utilizaram o algoritmo de *Louvain* [Blondel et al. 2008] para determinar comunidades e métricas (e.g., *variation of information*) para compará-las.

¹<http://lavajato.mpf.mp.br/perguntas-e-respostas>

No trabalho de Rostami and Mondani [2015], os autores utilizaram três bases de dados sobre uma mesma gangue de rua para construir redes e as compararam em relação a diferentes técnicas (visualização de redes, similaridade de redes e medidas de redes). Como resultado, os autores ressaltaram que as interpretações devem ser cuidadosas, já que dependendo da base e da medida utilizada, mesmos indivíduos podem possuir importâncias diferentes. Bright et al. [2012] e Morselli [2010] estudaram o cenário do narcotráfico. Morselli investigou o posicionamento de indivíduos em uma rede para verificar quais deles estão em posições vulneráveis e/ou estratégicas. Já Bright et al. analisaram a viabilidade da utilização de comentários de jurados como base de dados para aplicar técnicas de análise de redes. Em ambos os trabalhos, foram utilizadas duas medidas de centralidades: *degree* e *betweenness*.

Medidas de centralidade quantificam a importância de nós em redes sob diferentes aspectos. O grau ponderado, por exemplo, corresponde à soma dos pesos das arestas incidentes em um vértice v [Beveridge and Shan 2016]. Já a medida *betweenness*, representa o número de caminhos mínimos que passam por um vértice [Otte and Rousseau 2002]. A medida *closeness* corresponde a distância de um vértice v em relação a todos os outros vértices da rede [Freeman 1978]. As medidas *eigenvector* e *pagerank* consideram que um nó central é aquele que está conectado a outros nós também centrais [Bonacich 2007].

Nos trabalhos de Mancuso [2014] e Morselli and Savoie-Gargiso [2014] são estudadas redes criminais ligadas ao tráfico humano. Em Mancuso, são aplicadas técnicas de análise de redes para determinar o papel de senhoras (*madams*) em um esquema de tráfico sexual nigeriano. Para determinar a influência das *madams*, mulheres que escravizam meninas, foi utilizada a medida *betweenness*. Os resultados confirmaram a importância das *madams* nessas redes, embora nem todas elas sejam igualmente centrais. No trabalho de Morselli and Savoie-Gargiso, os autores investigaram aspectos como a coerção, controle e colaboração em uma rede de prostituição do Canadá. Para tanto, são utilizadas as medidas *degree*, *betweenness* e *clustering coefficient*. Como resultado, os autores destacaram que embora prostitutas sejam alvo de coerção e controle, algumas delas possuem posições privilegiadas nas redes, atuando como colaboradas do esquema de tráfico.

Em Liu et al. [2012] e Snijders and Baerveldt [2003], são analisadas redes criminais compostas por adolescentes delinquentes. Liu et al. propuseram um modelo baseado no equilíbrio de Nash, para identificar a personagem chave dessas redes criminais. Como resultado, foi verificado que os adolescentes que exercem um papel chave nas redes são homens, possuem pais com pouco nível de educação, estão pouco ligados a religiões e se sentem socialmente excluídos. Adicionalmente, estes jovens ainda que não sejam criminalmente ativos, podem exercer um papel chave na rede em termos da medida *betweenness*. Snijders and Baerveldt aplicaram métodos de Monte Carlo e redes de Markov para modelar a evolução de redes. Como estudo de caso, foi utilizada uma rede representando adolescentes com comportamento delincente. Os resultados evidenciaram que há um efeito de similaridade entre o comportamento delincente e a evolução das amizades: um alto grau de delinquência implica em um efeito positivo em relação a formação de conexões (amizades), mas também na sua dissolução.

Este trabalho diferencia-se dos demais por basear suas análises em medidas de centralidade selecionadas a partir da entropia de *Shannon*. Por meio da entropia, é possível verificar quais medidas são mais discriminativas para identificar indivíduos in-

fluentes em depoimentos. A entropia de *Shannon* pode ser definida como uma medida para quantificar a incerteza de uma distribuição p [Lesne 2014]. O valor da entropia é máximo quando a distribuição p é uniforme. Seja x uma variável aleatória, com valores pertencentes a um conjunto finito χ , a entropia normalizada de x é formulada como:

$$H(\chi) = -\frac{\sum_{x \in \chi} p(x) \log p(x)}{\log |\chi|}. \quad (1)$$

3. Metodologia Proposta

A metodologia utilizada neste trabalho é composta pelas seguintes etapas: (i) transcrição das delações premiadas para arquivos de textos; (ii) modelagem das redes complexas baseadas nas ocorrências conjuntas das entidades em trechos das delações; (iii) detecção de comunidades nas redes complexas; e (iv) identificação de entidades influentes, a partir de medidas de centralidade. Nas subseções que seguem, detalhamos estas etapas.

3.1. Base de Dados

A base de dados utilizada neste trabalho é composta por cinco delações premiadas homologadas pelo Supremo Tribunal Federal entre os anos de 2014 e 2017. Os depoentes das delações analisadas são:

1. Paulo Roberto Costa (<https://goo.gl/x0vo7z>);
2. Alberto Youssef (<https://goo.gl/MSLpBS>);
3. Nestor Cerveró (<https://goo.gl/cocruK>);
4. Delcídio do Amaral (<https://goo.gl/ZjIV1I>);
5. Cláudio Melo Filho (<https://goo.gl/pUvNRh>).

As delações premiadas são disponibilizadas à população em documentos no formato de arquivos PDF. Como os documentos foram digitalizados, há presença de ruídos que podem ocasionar a deformação de caracteres, dificultando assim qualquer análise sobre os dados. Embora existam ferramentas online que realizam esta tarefa, elas não apresentaram bom desempenho em razão da qualidade da digitalização das delações. Desta forma, a conversão foi feita manualmente.

3.2. Modelagem das Redes Complexas

Modelamos as redes complexas das delações premiadas como grafos não direcionados $G(V, E)$, onde V corresponde ao conjunto de vértices $\{v_1, v_2, \dots, v_n\}$ e E corresponde ao conjunto de arestas $\{e_1, e_2, \dots, e_k\}$ do grafo. Os vértices referem-se às pessoas físicas que tiveram seus nomes citados nas delações. As arestas referem-se à ocorrência conjunta dessas pessoas em trechos das delações, e seu peso é determinado pela quantidade dessas ocorrências conjuntas. Como os delatores estão sempre relacionados com as outras entidades de seus depoimentos, estes são, por definição, entidades relevantes sempre identificadas pelas análises de redes complexas. Por esta razão, o delator é excluído do conjunto de vértices de sua própria delação, podendo, entretanto, aparecer na delação de outrem. Após a modelagem das redes complexas, observamos que haviam alguns componentes desconexos no grafo que representavam entidades isoladas, isto é, que não ocorreram conjuntamente com nenhuma outra, ou entidades que ocorreram com outras, mas ambas não se relacionaram com nenhuma outra entidade da rede. Os componentes desconexos foram eliminados por não representarem informação pertinente.

3.3. Detecção de Comunidades

Após a modelagem e construção das redes complexas das delações, aplicamos o algoritmo de *Louvain* para identificar comunidades nessas redes. Com isso, buscamos verificar se os agrupamentos encontrados pelo algoritmo, refletiam algum tipo organização social.

3.4. Identificação de Entidades Influentes

Para avaliar a qualidade, em termos de capacidade discriminante, de cada métrica de centralidade, propomos o uso da entropia normalizada de *Shannon*. Neste caso, cada métrica é transformada em uma distribuição de probabilidade discreta (pdf) tal que: (1) cada classe da pdf é um vértice; (2) o valor da métrica é normalizado pela soma total. Desta forma, a soma de todas as classes da pdf é sempre um (requisito para uma pdf). A seguir, utilizamos a entropia normalizada de *Shannon* para determinar as métricas mais discriminantes. Neste caso, quanto maior a entropia de uma métrica, mais uniforme é a sua distribuição junto aos vértices da rede. Portanto, quanto menor a entropia, mais discriminante é a métrica em questão, pois entropia igual a **zero** refere-se ao caso determinístico, enquanto entropia igual a **um** refere-se a distribuição uniforme.

Tabela 1. Entropia média das medidas de centralidade.

Medida de Centralidade	Valor da Entropia
<i>Betweenness</i>	0.739
<i>Closeness</i>	0.996
<i>Eigenvector</i>	0.758
<i>Pagerank</i>	0.916
Grau Ponderado	0.862

Como é possível observar na tabela 1, as medidas de centralidade com menor entropia, considerando todas as redes complexas, foram *betweenness*, *eigenvector* e o grau ponderado respectivamente. Assim, utilizamos estas três medidas para determinar entidades influentes. Embora as demais medidas tenham apresentado entropia elevada, elas descrevem propriedades interessantes sobre as redes. Por exemplo, o elevado valor da entropia para a medida *closeness* indica que os vértices estão interligados a todos os outros por um número pequeno de caminhos, o que sugere que é necessária uma quantidade de tempo pequena para uma informação se espalhar de um nó para todos os outros. Entretanto, a alta entropia da medida *closeness* significa que todos os nós possuem valores muito próximos, sendo esta, uma métrica pouco discriminante.

4. Resultados

Esta seção apresenta a análise da influência de entidades em uma rede complexa gerada a partir da união das cinco redes provenientes de delações. Adicionalmente, analisamos alguns aspectos das comunidades identificadas nessas redes. Ressaltamos que os resultados encontrados são fruto da análise estrutural de redes provenientes de depoimentos, isto é, as informações encontradas refletem tão somente a narrativa dos depoimentos.

4.1. Análise das Entidades Influentes Identificadas

Para identificar as entidades mais influentes, definimos o *score* $r(v)$, normalizado, que pondera as três métricas de centralidade mais discriminantes de cada vértice v :

$$r(v) = \sqrt[3]{BE \times EV \times GP} \quad (2)$$

onde BE é o *betweenness*, EV o *eigenvector* e GP o grau ponderado (todos normalizados).

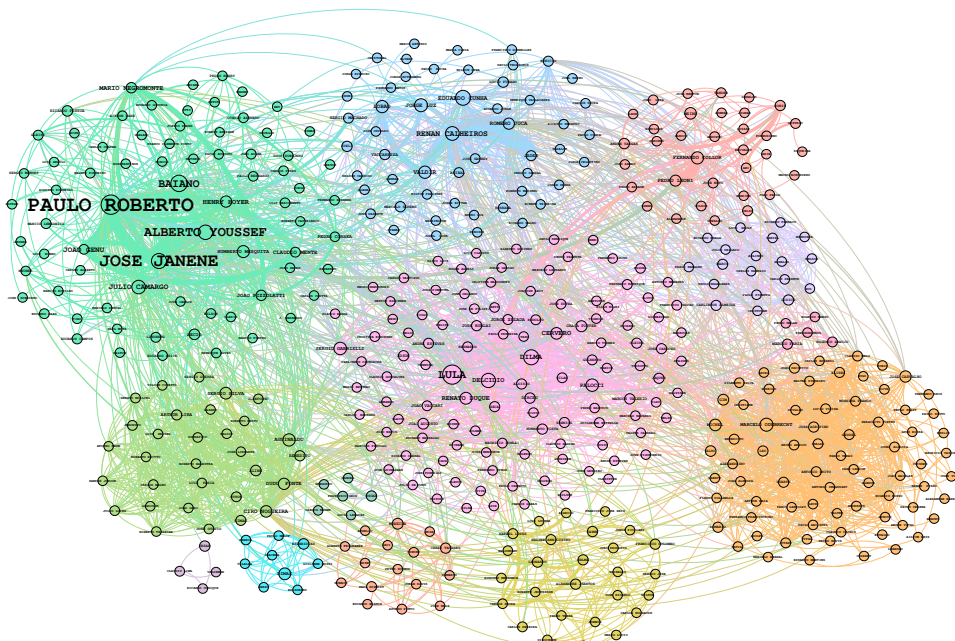


Figura 1. Rede Complexa gerada a partir da união das redes complexas das cinco delações premiadas. O tamanho dos vértices representam seus valores de *betweenness*, enquanto o tamanho de seus rótulos representam os valores da medida *eigenvector*.

Como as cinco delações foram narradas por depoentes que ocupavam posições distintas (doleiros, presidentes de empresas e políticos), a partir da união dessas delações é possível analisar o esquema investigado sob diferentes perspectivas, além de reduzir o viés dos depoentes sobre os fatos relacionados. A rede complexa da figura 1 possui um total de 468 vértices e 2749 arestas. A análise estrutural dessa rede indica que há 15 entidades-chave nas cinco delações consideradas. A tabela 2 mostra a situação atual destas quinze entidades mais influentes que identificamos. Dentre elas, somente uma não foi denunciada pelas autoridades (Jose Janene), uma vez que já faleceu. As demais entidades, estão sob investigação ou já foram condenadas. Desta forma, é possível verificar que nossa abordagem de redes complexas encontrou entidades que de fato exerciam influência no esquema investigado.

Em relação a medida *betweenness*, as entidades mais influentes são Paulo Roberto Costa, Luiz Inácio Lula da Silva e Fernando Baiano, respectivamente. Estas entidades representam vértices que atuam como pontes de comunicação entre diferentes partes da rede, isto é, constituem entidades com grande controle na rede, uma vez que muitas informações passam por elas. As entidades com maior valor para a medida *eigenvector* são Jose Janene, Paulo Roberto Costa e Alberto Youssef, respectivamente. Estas entidades

representam vértices que possuem conexões com entidades importantes, o que implica no aumento das suas próprias influências.

Tabela 2. Situação atual das entidades influentes identificadas e posição segundo as medidas de centralidade *betweenness* (BE), *eigenvector* (EV) e grau ponderado (GP).

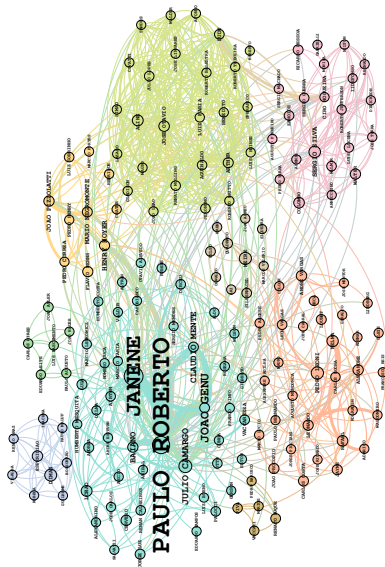
Entidades Influentes	Situação	Posição (ranking)			Score
		BE	EV	GP	
Paulo Roberto Costa	condenado	1.000 (1 ^o)	0.886 (2 ^o)	1.000 (1 ^o)	0.960
Jose Janene	falecido	0.562 (4 ^o)	1.000 (1 ^o)	0.821 (2 ^o)	0.773
Alberto Youssef	condenado	0.500 (7 ^o)	0.690 (3 ^o)	0.520 (5 ^o)	0.564
Luiz Inácio Lula da Silva	réu	0.975 (2 ^o)	0.303 (8 ^o)	0.590 (4 ^o)	0.558
Fernando Baiano	condenado	0.700 (3 ^o)	0.401 (5 ^o)	0.473 (7 ^o)	0.510
Renan Calheiros	réu	0.429 (9 ^o)	0.259 (10 ^o)	0.807 (3 ^o)	0.447
Dilma Rousseff	investigada	0.557 (5 ^o)	0.154 (18 ^o)	0.448 (9 ^o)	0.337
Nestor Cerveró	condenado	0.210 (23 ^o)	0.284 (9 ^o)	0.509 (6 ^o)	0.312
Júlio Camargo	condenado	0.408 (10 ^o)	0.229 (13 ^o)	0.289 (24 ^o)	0.300
Delcídio do Amaral	condenado	0.479 (8 ^o)	0.115 (27 ^o)	0.389 (10 ^o)	0.278
Renato Duque	condenado	0.240 (20 ^o)	0.246 (12 ^o)	0.331 (18 ^o)	0.269
Henry Hoyer	investigado	0.178 (28 ^o)	0.314 (7 ^o)	0.309 (19 ^o)	0.258
Valdir Raupp	réu	0.280 (15 ^o)	0.178 (15 ^o)	0.295 (22 ^o)	0.245
Ciro Nogueira	réu	0.206 (26 ^o)	0.175 (16 ^o)	0.373 (11 ^o)	0.237
Romera Jucá	investigado	0.169 (30 ^o)	0.164 (17 ^o)	0.470 (9 ^o)	0.235

Quanto ao grau ponderado, as entidades com maior influência são Paulo Roberto Costa, Jose Janene e Renan Calheiros, o que implica que estes vértices além de possuírem muitas conexões, detêm uma relação forte com sua vizinhança. Por fim, as entidades que possuem maior influência relativa às três medidas de centralidade são Paulo Roberto Costa, Jose Janene e Alberto Youssef. Estes vértices representam entidades que se relacionam fortemente com sua vizinhança, além de estarem interligados com entidades também influentes. Adicionalmente, caso esses vértices fossem retirados, a comunicação entre diferentes partes da rede seria prejudicada.

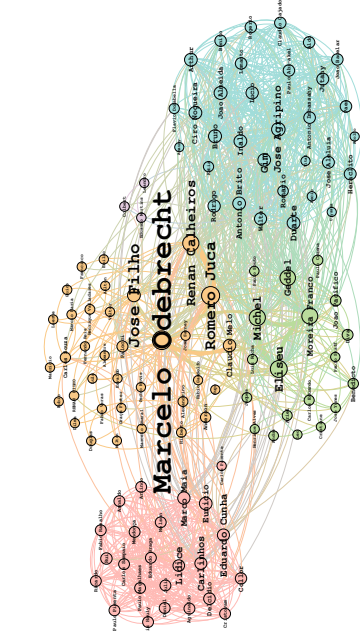
4.2. Análise das Comunidades Identificadas

A figura 2 apresenta as redes complexas das delações que compõe nossa base de dados. Os agrupamentos de vértices de mesma cor representam as comunidades identificadas. De forma geral, as comunidades presentes nas redes refletem esquemas de investigados ou integrantes de uma mesmo partido político ligados a fatos que os delatores descreveram.

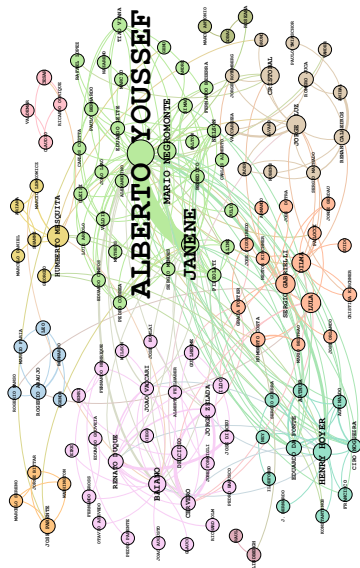
Quanto aos agrupamentos de entidades pertencentes a partidos políticos, estes podem ser observados na comunidade verde das delações de Paulo Roberto Costa, Cláudio Melo Filho, Nestor Cerveró e nas comunidades amarela e verde da delação de Alberto Youssef. Os partidos implicados nessas comunidades são PMDB (Partido do Movimento Democrático Brasileiro), PP (Partido Progressista) e PT (Partido dos Trabalhadores).



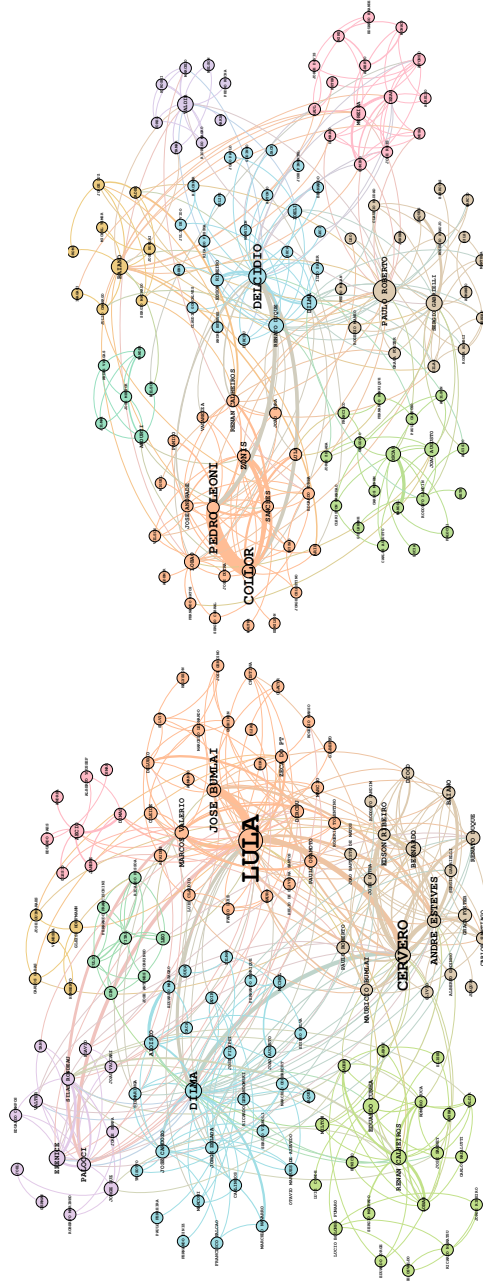
(c) Delação de Alberto Youssef.



(b) Delação de Cláudio Melo Filho.



(a) Delação de Paulo Roberto Costa.



(e) Delação de Nestor Cerveró.

(d) Delação de Delcídio do Amaral.

Figura 2. Redes Complexas das delações premiadas de Paulo Roberto Costa, Alberto Youssef, Nestor Cerveró, Delcídio do Amaral e Cláudio Melo Filho.

As demais comunidades das redes da figura 2 são referentes a esquemas de repasses de valores a políticos. Por exemplo, a comunidade laranja da delação de Alberto Youssef concentra entidades que participaram de um esquema ilícito relacionado a uma rede de postos de gasolina. Além de entidades que prestavam serviços para o declarante. É importante ressaltar que mesmo que uma entidade esteja presente em uma dessas comunidades de esquemas de repasses de valores, ela pode não ter nenhuma relação com tal esquema. Isso pode ser reflexo da forma como o algoritmo de *Louvain* agrupa os vértices.

Na figura 2 o tamanho do vértice representa seu valor de *betweenness* e o tamanho de seu rótulo equivale ao valor de *eigenvector*. É interessante observar que na rede de Paulo Roberto Costa, a entidade que com maior influência é Alberto Youssef. Já na rede do próprio Alberto Youssef, a entidade com maior influência é Paulo Roberto Costa. Isto sugere uma forte ligação entre essas duas entidades. Adicionalmente, ambas estão conectadas fortemente com Jose Janene, o que se reflete nos resultados da subseção anterior.

Por meio das redes complexas que modelam as delações, é possível verificar que vértices que se referem à mesma entidade possuem influências variadas. Como exemplo, destacamos a entidade Luiz Inácio Lula de Silva e Renan Calheiros que apareceram em todas as redes, porém com valores bem diferentes para as medidas de centralidade. Tal fato, foi também observado por Rostami and Mondani [2015].

5. Conclusões e Trabalhos Futuros

Neste trabalho apresentamos uma abordagem de redes complexas para identificar entidades influentes em depoimentos por meio de medidas de centralidade. Desta forma, autoridades policiais podem direcionar suas investigações para aqueles indivíduos que exercem papel chave nessas redes. Como estudo de caso, apresentamos redes complexas construídas a partir de delações premiadas provenientes da Operação Lava Jato. Adicionalmente, identificamos e analisamos comunidades nessas redes para verificar se elas refletiam algum tipo de organização social particular. Nossos resultados encontrados apresentam coerência com o histórico de pessoas denunciadas/indiciadas pelo Ministério Público e pela Procuradoria Geral da República.

Dentre os desafios deste trabalho, destacamos o processo de transcrição das delações premiadas, além de erros de grafia presentes nas delações. Além disso, tivemos que tratar os dados de forma cuidadosa, uma vez que envolvem nomes de figuras públicas. Como limitação, a análise estrutural das redes complexas modeladas refletem o testemunho dos declarantes. Portanto, a validade das análises estruturais e sociais apresentadas são dependentes da veracidade das informações contidas nos depoimentos. Além disso, há o viés, intencional ou não, do declarante. Neste trabalho pudemos constatar que a análise conjunta de vários depoimentos em uma única rede tende a mitigar o viés dos declarantes. Entretanto, a detecção de inconsistências, falácias e viés não foram alvo deste trabalho.

Como trabalho futuro pretendemos investigar novas técnicas para determinar vértices influentes em grafos, além de examinar técnicas para verificar a importância dos subgrupos de vértices que formam comunidades, para assim, aprofundar nossa análise em redes criminais. Outro trabalho futuro consiste na modelagem e análise de redes complexas para identificação de inconsistências, falácias e viés em depoimentos.

Referências

- Beveridge, A. and Shan, J. (2016). Network of thrones. *Math Horizons*, 23(4):18–22.
- Blondel, V., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social networks*, 29(4):555–564.
- Brandão, M. and Moro, M. (2016). Social professional networks: A survey and taxonomy. *Computer Communications*, 100.
- Bright, D., Hughes, C., and Chalmers, J. (2012). Illuminating dark networks: A social network analysis of an Australian drug trafficking syndicate. *Crime, Law and Social Change*, 57(2):151–176.
- Calderoni, F. and Piccardi, C. (2014). Uncovering the structure of criminal organizations by community analysis: The infinito network. In *Proc. of 10th Int'l Conf. on Signal-Image Technology and Internet-Based Systems*, pages 301–308.
- Campana, P. (2016). Explaining criminal networks: Strategies and potential pitfalls. *Methodological Innovations*, 9:1–10.
- Freeman, L. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Kossinets, G. and Watts, D. (2009). Origins of homophily in an evolving social network. *American journal of sociology*, 115(2):405–450.
- Lesne, A. (2014). Shannon entropy: A rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics. *Mathematical Structures in Computer Science*, 24(03):240–311.
- Liu, X., Patacchini, E., Zenou, Y., and Lee, L.-F. (2012). Criminal networks: Who is the key player?
- Mancuso, M. (2014). Not all madams have a central role: analysis of a Nigerian sex trafficking network. *Trends in Organized Crime*, 17(1-2):66–88.
- Masys, A. (2013). Networks and network analysis for defence and security. In *Proc. of the Int'l Conf. on Advances in Social Networks Analysis and Mining*, pages 1479–1480.
- Morselli, C. (2010). Assessing vulnerable and strategic positions in a criminal network. *Journal of Contemporary Criminal Justice*, 26(4):382–392.
- Morselli, C. (2013). *Crime and Networks*. Routledge.
- Morselli, C. and Savoie-Gargiso, I. (2014). Coercion, control, and cooperation in a prostitution ring. *The ANNALS of the American Academy of Political and Social Science*, 653(1):247–265.
- Otte, E. and Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of information Science*, 28(6):441–453.
- Rostami, A. and Mondani, H. (2015). The complexity of crime network data: A case study of its consequences for crime control and the study of networks. *PloS One*, 10(3):e0119309.
- Snijders, T. and Baerveldt, C. (2003). A multilevel network study of the effects of delinquent behavior on friendship evolution. *Journal of mathematical sociology*, 27(2-3):123–151.